

RESEARCH

Open Access



Robust beamforming designs for multiuser MISO downlink with per-antenna power constraints

Feng Wang¹, Yongwei Huang², Xin Wang^{1*} and Yu Zhu¹

Abstract

This paper studies the robust beamforming designs for a multiuser multiple-input single-output (MISO) downlink system. Different from the conventional sum-power constraint across all transmit antennas, we consider individual power constraints per antenna at the base station. Assuming that the channel uncertainty is bounded by a spherical region, we develop the optimal robust designs to maximize the minimum worst-case signal-to-interference-plus-noise ratio (SINR) among all users. Specifically, we show that the optimal max-min SINR beamformers can be obtained by solving a sequence of “dual” min-max power problems. Relying on the S-procedure and the linear matrix inequality representation for the cone of Lorentz-positive maps (LPMs), respectively, two designs, which are referred to as Robust-SP and Robust-LPM, are proposed to efficiently solve such min-max power problems. Building on either Robust-SP or Robust LPM, a bisection search algorithm is then developed to find the robust max-min SINR beamformers with guaranteed global optimality and geometrically fast convergence speed. Using the max-min SINR solutions as a cornerstone, we further put forth the optimal robust design for the worst-case weighted sum-rate (WSR) maximization. By formulating the worst-case WSR maximization problem into a monotonic program (MP), we develop a polyblock outer approximation algorithm to obtain the globally optimal solution. Numerical results are presented to demonstrate the merits of the proposed robust beamforming designs.

Keywords: Lorentz-positive maps; MISO robust beamforming; Monotonic program; Per-antenna power constraints; Semidefinite program

1 Introduction

For a multiuser wireless downlink with multiple antennas at the base station (BS), transmit beamforming is a low-complexity solution to provide spatial diversity and to reduce co-channel interference [1]. Under the assumption of perfect channel state information (CSI), the optimal and suboptimal beamforming designs were developed to maximize the minimum user signal-to-interference-plus-noise ratio (SINR) [1–4] or to maximize the weighted sum of user rates [5–7]. However, perfect CSI is usually unavailable in practical systems due to many practical factors, such as inaccurate channel estimation, quantization of CSI, erroneous or outdated feedback, and time delays

or frequency offsets between the reciprocal channels. For these reasons, robust linear and nonlinear beamforming (precoding) designs to combat against channel uncertainty have received intensive research interests for different multi-user multiple-input multiple-output (MIMO) (single-/multi-cell downlink or relay) systems [8–19]. Generally, robust beamforming is addressed by either a stochastic or a worst-case approach. In the stochastic approach, the CSI uncertainty is often modeled as a random variable subject to a known probability distribution, and the robustness can be provided based on optimizing the average or the outage performance such as sum mean-square-error (MSE) [8] and weighted-sum-rate (WSR) [9, 10]. On the other hand, the worst-case approach assumes that the CSI lies in a bounded uncertainty region. In this case, the robustness is achieved by optimizing the system under the worst-case channel condition and

*Correspondence: xwang11@fudan.edu.cn; eeyw@ece.ust.hk

¹The Key Laboratory for Information Science of Electromagnetic Waves (MoE), the Department of Communication Science and Engineering, Fudan University, 220 Han Dan Road, Shanghai, China

Full list of author information is available at the end of the article

it usually leads to a min-max or max-min problem formulation. With this approach, robust linear beamforming designs were developed for max-min SINR [11–16] or WSR maximization [17–19].

All the existing beamforming designs in [1–19] were developed under a sum-power constraint across all transmit antennas. However, in physical implementation of a multi-antenna BS, each antenna usually has its own power amplifier in the analog front-end, and it is individually limited by the linearity of this amplifier [20]. Hence, instead of the sum-power constraint, this leads to power constraints imposed on a per-antenna basis. Under such more realistic power constraints, the transmitter optimization was addressed in [20], where the elegant uplink-downlink duality under a sum-power constraint was extended to downlink problems with per-antenna power constraints. The robust sum-MSE minimizations with per BS antenna and per BS power constraints were investigated in [21] through the downlink-uplink duality. By establishing the MSE downlink-interference duality, the authors in [22] extended the work of [21] to solve the weighted sum-MSE minimization and min-max MSE problems under general power constraints for multiuser MIMO systems. Downlink beamforming designs under per-antenna power constraints were also addressed in some other works. Moore-Penrose zero forcing (ZF)-based beamforming design was derived to maximize the minimum user rate in [23]. Aiming to minimize the stochastic or worst-case sum-MSE, a nonlinear Tomlinson-Harashima precoding (THP) design was investigated in [24] under per-antenna power constraints. Based on block coordinate ascent and signomial programming methods, beamforming designs were put forth to maximize the WSR for a multiuser downlink with power constraints per antenna groups in [25]. For a large antenna array at the BS, an achievable rate for single-user multiple-input single-output (MISO) beamforming under per-antenna constant-envelope constraints was derived in [26]. With per-antenna array power constraints, the surrogate duality of the max-min beamforming was investigated in [27].

In this paper, we investigate the robust transmit beamforming designs for a MISO downlink with norm-bounded uncertain CSI at the BS. Under the per-antenna power constraints, the robust designs aim to maximize the minimum worst-case SINR among all users, or to maximize the worst-case WSR. The main contributions are summarized as follows:

- This paper shows that the robust max-min SINR beamformers can be obtained by solving a sequence of “dual” robust min-max power problems. Building on the solutions to these “dual” problems, an efficient bisection search algorithm can thus be applied to find the optimal max-min SINR solution.
- Two robust designs (i.e., Robust-SP and Robust-LPM) are developed to solve the “dual” robust min-max power problem. This problem is well-known nonconvex due to the infinitely many SINR constraints. For Robust-SP, we use the S-procedure to convert the problem into a rank-constrained semidefinite program (SDP), and then apply the SDP relaxation technique to find its (near-)optimal solution. Like [15], we give a computable CSI uncertainty bound which ensures the tightness of the SDP relaxation. For Robust-LPM, we consider a slightly conservative problem reformulation. Relying on a linear matrix inequality (LMI) representation for the cone of Lorentz-positive maps (LPMs), the new problem is shown to be equivalently transformed into a convex SDP which can be efficiently solved with guaranteed global optimality.
- We further put forth the robust beamforming for the worst-case WSR maximization and formulate this problem into a monotonic program (MP). Relying on solving a sequence of robust max-min SINR problems, a polyblock outer approximation algorithm is developed to find the globally optimal solution.

The remainder of this paper is organized as follows. Section 2 presents the system model. Section 3 introduces the proposed approach to the optimal max-min SINR beamforming design with perfect CSI. Section 4 gives the problem formulation of the optimal robust max-min SINR beamforming, and then develops Robust-SP and Robust-LPM to solve its “dual” min-max power problem. The robust design for the worst-case WSR maximization is presented in Section 5. Numerical results are provided in Section 6, followed by the conclusion in Section 7.

Notations: Uppercase and lowercase boldface letters denote matrices and vectors, respectively. All vectors are column vectors. \mathbb{R}^n and \mathbb{C}^n are the n -dimensional spaces of real and complex vectors, respectively. $\mathbb{E}[\cdot]$ denotes the expectation, $(\cdot)^T$ is the transpose, $(\cdot)^H$ is the Hermitian transpose, $|\cdot|$ is the absolute value, and $\|\cdot\|$ is the Euclidean norm of a vector. $\mathbf{0}$ and $\mathbf{1}$ are all-zero and all-one vectors of appropriate dimension, respectively. $\mathcal{R}(\cdot)$ and $\mathcal{I}(\cdot)$ denote the real part and imaginary part of the argument, respectively. The letter j represents the imaginary unit (i.e., $j = \sqrt{-1}$), while the letter i often serves as an index in this paper. \mathbf{I}_n denotes an $n \times n$ identity matrix. $\mathbf{A} \succeq \mathbf{0}$ means that \mathbf{A} is Hermitian and positive semi-definite, $[\mathbf{A}]_{n,n}$ is the (n, n) -th entry of \mathbf{A} . $\text{tr}(\mathbf{A})$ and $\text{rank}(\mathbf{A})$ are the trace and rank of \mathbf{A} , respectively. \otimes denotes the Kronecker product. A circular symmetric complex Gaussian (CSCG) random vector \mathbf{a} with mean $\bar{\mathbf{a}}$ and covariance matrix Σ is denoted as $\mathbf{a} \sim \mathcal{CN}(\bar{\mathbf{a}}, \Sigma)$. $\mathcal{A} \cup \mathcal{B}$ and $\mathcal{A} \cap \mathcal{B}$ denote the union and intersection of set \mathcal{A} and set \mathcal{B} , respectively.

$\mathcal{A} \setminus \mathcal{B}$ denotes the part of \mathcal{A} that does not intersect with \mathcal{B} . $\mathcal{A} \subset \mathcal{B}$ denotes that \mathcal{A} is contained in \mathcal{B} . Finally, the element-wise (strictly) inequality for \mathbf{x} , \mathbf{y} is denoted as $\mathbf{x} \leq \mathbf{y}$ ($\mathbf{x} < \mathbf{y}$).

2 Modeling preliminaries

Consider a wireless multiuser MISO downlink system where the BS equipped with M antennas simultaneously transmits K independent signals to K single-antenna users, as shown in Fig. 1. The baseband equivalent representation of the received signal at user k is expressed as:

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \forall k, \quad (1)$$

where $\mathbf{h}_k \in \mathbb{C}^M$ is the channel vector from the BS to user k , $\mathbf{x} \in \mathbb{C}^M$ is the transmitted signal at the BS, $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive white complex Gaussian noise. Assuming that a linear beamforming strategy is employed at the BS, the transmitted signal \mathbf{x} at the BS is then given by

$$\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k, \quad (2)$$

where $\mathbf{w}_k \in \mathbb{C}^M$ and $s_k \in \mathbb{C}$ are the beamforming vector and the information signal for user k , respectively. Without loss of generality (w.l.o.g.), we assume that $s_k, \forall k$, has a unit power, i.e., $\mathbb{E}[|s_k|^2] = 1$.

Suppose that each transmit antenna has its own power constraint $P_m, \forall m$. Given the beamformers $\mathbf{w}_k, \forall k$, such per-antenna power constraints can be expressed as

$$\left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{m,m} \leq P_m, \forall m. \quad (3)$$

Define $\mathbf{W} := [\mathbf{w}_1, \dots, \mathbf{w}_K]$, then the received SINR at user k in terms of \mathbf{W} is given by

$$\text{SINR}_k(\mathbf{W}) = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{l=1, l \neq k}^K |\mathbf{h}_k^H \mathbf{w}_l|^2 + \sigma_k^2}, \forall k. \quad (4)$$

3 Max-min SINR beamforming with perfect CSI

As a good starting point, we first develop an efficient approach to find the optimal beamformers that maximize the minimum SINR among all users under the per-antenna power constraints (3), when the perfect CSI is available at the BS.

Based on the SINRs in (4), we consider the following classic max-min SINR balancing problem:

$$\begin{aligned} & \max_{\mathbf{W}} \min_{1 \leq k \leq K} \frac{\text{SINR}_k(\mathbf{W})}{\gamma_k} \\ & \text{s. t.} \quad \left[\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H \right]_{m,m} \leq P_m, \forall m, \end{aligned} \quad (5)$$

where $\gamma_k > 0$ is the prescribed SINR target for user k . Define the $M \times M$ square matrix \mathbf{D}_m with

$$[\mathbf{D}_m]_{i,l} := \begin{cases} 1, & i = l = m, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

The per-antenna power constraints of (5) are then rewritten as

$$\sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq P_m, \forall m, \quad (7)$$

which are convex quadratic inequality constraints with respect to \mathbf{w}_k .

Due to the nonconvexity of $\text{SINR}_k(\mathbf{W})$, the max-min SINR problem (5) is non-convex, and hence difficult to solve. Yet, we next show that it can be solved by alternatively solving the tractable “dual” per-antenna power balancing problems.

For a given scalar $\lambda > 0$, as shown in [28], by introducing an auxiliary variable α , we formulate the following per-antenna power balancing problem:

$$\begin{aligned} & \alpha^*(\lambda) = \min_{\mathbf{W}, \alpha} \alpha \\ & \text{s. t.} \quad \frac{\sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k}{P_m} \leq \alpha, \forall m, \\ & \quad \frac{\text{SINR}_k(\mathbf{W})}{\gamma_k} \geq \lambda, \forall k, \end{aligned} \quad (8)$$

where the SINR constraints in (8) can be rewritten as:

$$\left(1 + \frac{1}{\lambda \gamma_k}\right) |\mathbf{h}_k^H \mathbf{w}_k|^2 \geq \left\| \begin{bmatrix} \mathbf{h}_k^H \mathbf{W} \\ \sigma_k \end{bmatrix} \right\|^2, \forall k. \quad (9)$$

Observe that an arbitrary phase rotation can be added to the beamformers without affecting the SINRs. Hence, we choose phases such that $\mathbf{h}_k^H \mathbf{w}_k, \forall k$, are real and nonnegative. By taking the square root operation for both sides of (9), the constraints then become convex second-order cone (SOC) (as known as Lorentz cone) constraints [29].

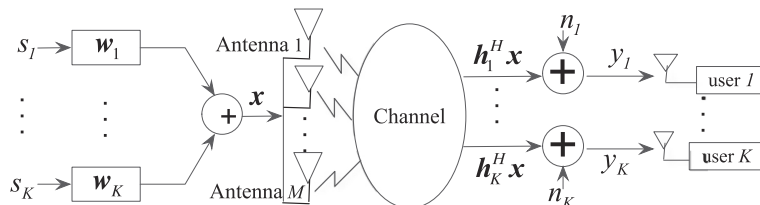


Fig. 1 The multiuser MISO downlink model

As a result, (8) can be reformulated into the following convex SOCP problem:

$$\begin{aligned} \alpha^*(\lambda) = \min_{\mathbf{W}, \alpha} \quad & \alpha \\ \text{s. t.} \quad & \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq \alpha P_m, \forall m, \\ & \sqrt{1 + \frac{1}{\lambda \gamma_k}} \mathbf{h}_k^H \mathbf{w}_k \geq \left\| \begin{bmatrix} \mathbf{h}_k^H \mathbf{W} \\ \sigma_k \end{bmatrix} \right\|, \forall k, \end{aligned} \quad (10)$$

which can be efficiently solved by interior-point methods with given parameter λ [29]. Note that, with the given γ_k , $\forall k$, the problem (10) may be infeasible for large positive λ , since that the SINR constraints cannot be satisfied. In this case, we define the optimal value $\alpha^*(\lambda) := +\infty$. When (10) is feasible, we define $\mathbf{W}^*(\lambda)$ as the optimal solution to (10) for a given λ . It can be shown that the function $\alpha^*(\lambda)$ obeys the following two properties¹:

- *Monotonicity*: If $0 < \lambda_1 \leq \lambda_2$, then $\alpha^*(\lambda_1) \leq \alpha^*(\lambda_2)$;
- *Optimum condition*: If it holds $\alpha^*(\lambda) = 1$, then λ and the corresponding $\mathbf{W}^*(\lambda)$ for (10) are the optimal value and the optimal solution for (5), respectively.

The two properties clearly indicate that the optimal solution of (5) can be obtained by solving the equation $\alpha^*(\lambda) = 1$. Building on the solution for (10), an efficient bisection search algorithm can be then implemented to find the max-min SINR beamformers for (5) with guaranteed global optimality and geometrically fast convergence speed [28].

It is worth mentioning that the max-min SINR problem (5) can be equivalently formulated as a quasi-concave problem:

$$\begin{aligned} \max_{\mathbf{W}} \quad & \lambda \\ \text{s. t.} \quad & \sqrt{1 + \frac{1}{\lambda \gamma_k}} \mathbf{h}_k^H \mathbf{w}_k \geq \left\| \begin{bmatrix} \mathbf{h}_k^H \mathbf{W} \\ \sigma_k \end{bmatrix} \right\|, \forall k, \\ & \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq P_m, \forall m, \end{aligned} \quad (11)$$

where the phases of $\mathbf{w}_k, \forall k$, are chosen such that $\mathbf{h}_k^H \mathbf{w}_k, \forall k$, are real and nonnegative. Clearly, (11) can be efficiently solved using a similar bisection search, where each step involves checking a convex feasibility problem with an updated λ :

$$\begin{aligned} \text{Find } \mathbf{W} \quad & \\ \text{s. t.} \quad & \sqrt{1 + \frac{1}{\lambda \gamma_k}} \mathbf{h}_k^H \mathbf{w}_k \geq \left\| \begin{bmatrix} \mathbf{h}_k^H \mathbf{W} \\ \sigma_k \end{bmatrix} \right\|, \forall k, \\ & \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq P_m, \forall m. \end{aligned} \quad (12)$$

We note that (10) and (12) correspond to a similar SOCP; hence, a similar computational effort is required to solve each instance of (10) and that of (12), practically using a software like CVX [29]. Also since that the similar bisection search methods are applied in the proposed scheme and the feasibility check approach, it is expected that both algorithms can achieve the optimal solution of (5) after the same iterations. While the proposed scheme can solve the max-min SINR problem with similar computational complexity compared to the feasibility check approach, it provides a new and important insight for the close relationship between the max-min SINR problem and the min-max power problem in beamforming designs. This allows us to freely choose a more tractable form as the corner stone to pursue the optimal beamforming designs under various important criteria such as rate maximization, MSE minimization, and bit-error-rate (BER) minimization, as detailed in Section 5.

4 Robust max-min SINR beamforming design

In this section, we generalize the approach in the perfect CSI case to the robust downlink beamforming designs where the BS has the bounded CSI uncertainty. We firstly formulate a robust max-min SINR beamforming problem and show that the optimal solution can be obtained by solving a sequence of its “dual” min-max power problems. An efficient bisection search algorithm is thus developed. Resorting to the S-procedure and the LMI representation for the cone of LPMs, respectively, Robust-SP and Robust-LPM are then proposed to solve these robust min-max power problems.

4.1 Robust max-min SINR beamforming problem

The downlink CSI uncertainty at the BS could be caused by estimation errors, feedback quantization, hardware deficiencies, and delays in CSI acquisition [11]. Motivated by these considerations, we assume the following additive error model:

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \delta_k, \forall k, \quad (13)$$

where $\hat{\mathbf{h}}_k$ is the estimated channel at the BS, and δ_k denotes the channel uncertainty. Similar to [11, 13, 15], we further assume that δ_k is bounded by a spherical region²:

$$\mathcal{H}_k := \left\{ \hat{\mathbf{h}}_k + \delta_k \mid \sqrt{\delta_k^H \delta_k} \leq \epsilon_k \right\}, \forall k, \quad (14)$$

where the real parameter $\epsilon_k > 0$ specifies the radius of \mathcal{H}_k .

Next, based on the CSI uncertainty region \mathcal{H}_k , define the worst-case SINR for user k in terms of \mathbf{W} as:

$$\widetilde{\text{SINR}}_k(\mathbf{W}) := \min_{\mathbf{h}_k \in \mathcal{H}_k} \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{l=1, l \neq k}^K |\mathbf{h}_k^H \mathbf{w}_l|^2 + \sigma_k^2}, \forall k. \quad (15)$$

We consider the robust designs aiming to maximize the minimum worst-case SINR among all users. By replacing

the SINRs in (5) with the worst-case SINRs in (15), we have the following robust max-min SINR balancing formulation:

$$\begin{aligned} \lambda^* = \max_{\mathbf{W}} \min_{1 \leq k \leq K} \frac{\widetilde{\text{SINR}}_k(\mathbf{W})}{\gamma_k} \\ \text{s. t. } \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq P_m, \forall m. \end{aligned} \quad (16)$$

Again the problem (16) is nonconvex. It is worth mentioning that, like the perfect CSI case, (16) can be converted into a quasi-concave problem with the aid of the S-procedure and rank-relaxation, and the (near-)optimal solution can be found by iteratively checking the feasibility of an SDP. Here, we show that (16) can be solved by alternatively solving a sequence of the min-max power problems. For a given $\lambda > 0$, we consider:

$$\begin{aligned} \tilde{\alpha}^*(\lambda) = \min_{\mathbf{W}} \max_{1 \leq m \leq M} \frac{\sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k}{P_m} \\ \text{s. t. } \frac{\widetilde{\text{SINR}}_k(\mathbf{W})}{\gamma_k} \geq \lambda, \forall k, \end{aligned} \quad (17)$$

where P_m serves as a “power target” for antenna m and $\tilde{\alpha}^*(\lambda)$ denotes the optimal value of problem (17). Again the problem (17) may be infeasible for large nonnegative λ . In this case, we define $\tilde{\alpha}^*(\lambda) := +\infty$. As with the perfect CSI case, we can establish the following property:

Lemma 1. $\tilde{\alpha}^*(\lambda)$ is a strictly increasing function for $\lambda > 0$.

Proof. Please see Appendix 1.

Let $\mathbf{W}^*(\lambda)$ denote the optimal beamforming matrix for (17) with the given $\lambda > 0$. Relying on the monotonicity of $\tilde{\alpha}^*(\lambda)$, we can further show the following close relationship between (16) and (17): \square

Lemma 2. If it holds $\tilde{\alpha}^*(\lambda) = 1$, then λ and the corresponding $\mathbf{W}^*(\lambda) = [\mathbf{w}_1^*(\lambda), \dots, \mathbf{w}_K^*(\lambda)]$ are the optimal value and the optimal solution for (16), respectively.

Proof. Please see Appendix 2. \square

Similar to the perfect CSI case, Lemma 2 indicates that the optimal solution to (16) can be obtained by solving the equation $\tilde{\alpha}^*(\lambda) = 1$. Therefore, a bisection search over λ can be then implemented to obtain the robust max-min SINR beamformers. We describe it in Algorithm 1 in detail.

For Algorithm 1, an appropriate interval of λ^* should be determined. Obviously, $\lambda_{\min} = 0$, a lower bound of the interval. An upper bound λ_{\max} is determined as follows. It is clear that,

Algorithm 1 for the max-min SINR beamforming problem (16)

- 1: **Initialize:** Select a tolerance level $\delta > 0$;
- 2: Set $\lambda_{\min} = 0, \lambda_{\max} = \max_k \left\{ \frac{(\|\hat{\mathbf{h}}_k\| + \|\epsilon_k\|)^2 (\sum_{m=1}^M P_m)}{\gamma_k \sigma_k^2} \right\}$;
- 3: **Repeat:** Let $\lambda = (\lambda_{\max} + \lambda_{\min})/2$;
- 4: Solving (17) to obtain $\tilde{\alpha}^*(\lambda)$ and $\mathbf{W}^*(\lambda)$;
- 5: If $\tilde{\alpha}^*(\lambda) \geq 1$, then $\lambda_{\max} = \lambda$;
- 6: Else if $\tilde{\alpha}^*(\lambda) < 1$, then $\lambda_{\min} = \lambda$;
- 7: **Until:** $(\lambda_{\max} - \lambda_{\min})/\lambda_{\max} < \delta$.
- 8: **Output:** $\lambda^* = \lambda$ and $\mathbf{W}^* = \mathbf{W}^*(\lambda)$.

$$\begin{aligned} \frac{\widetilde{\text{SINR}}(\mathbf{W})}{\gamma_k} &= \min_{\|\delta_k\| \leq \epsilon_k} \frac{|\hat{\mathbf{h}}_k + \delta_k|^H \mathbf{w}_k|^2}{\gamma_k \left(\sum_{l=1, l \neq k}^K |(\hat{\mathbf{h}}_l + \delta_l)^H \mathbf{w}_l|^2 + \sigma_k^2 \right)} \\ &\leq \frac{(\|\hat{\mathbf{h}}_k\| + \|\epsilon_k\|)^2 (\sum_{m=1}^M P_m)}{\gamma_k \sigma_k^2}, \quad \forall \mathbf{W}, \forall k. \end{aligned}$$

This implies that

$$\begin{aligned} \lambda^* &= \max_{\mathbf{W}} \min_{1 \leq k \leq K} \left\{ \frac{\widetilde{\text{SINR}}_k(\mathbf{W})}{\gamma_k} \right\} \\ &\leq \max_{1 \leq k \leq K} \left\{ \frac{(\|\hat{\mathbf{h}}_k\| + \|\epsilon_k\|)^2 (\sum_{m=1}^M P_m)}{\gamma_k \sigma_k^2} \right\}. \end{aligned} \quad (18)$$

With such λ_{\min} and λ_{\max} , the bisection search requires $\mathcal{O}(\log_2((\lambda_{\max} - \lambda_{\min})/\delta))$ iterations (e.g., see [29] page 146) to solve $\tilde{\alpha}^*(\lambda) = 1$ up to a desired accuracy level, due to the monotonicity of $\tilde{\alpha}^*(\lambda)$ with respect to $\lambda > 0$. Clearly, Algorithm 1 converges to the optimal solution geometrically fast.

Summarizing, we have the following proposition [29]:

Proposition 1. Algorithm 1 converges geometrically fast to the global optimal solution \mathbf{W}^* for (16).

To implement Algorithm 1, it is desirable to efficiently solve problem (17) with any given $\lambda > 0$. Introducing an auxiliary variable α , we rewrite (17) as:

$$\begin{aligned} \tilde{\alpha}^*(\lambda) &= \min_{\mathbf{W}, \alpha} \alpha \\ \text{s. t. } \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k &\leq \alpha P_m, \forall m, \\ \frac{\widetilde{\text{SINR}}_k(\mathbf{W})}{\gamma_k} &\geq \lambda, \forall k. \end{aligned} \quad (19)$$

Although (19) is still non-convex due to the infinitely many SINR constraints, two efficient designs will be provided in the next two subsections.

4.2 The Robust-SP design

The basic idea of Robust-SP is to firstly convert (19) into a rank-constrained SDP by applying the well-known S-procedure [29, 30], and then the converted problem is (near-)optimally solved via the SDP relaxation [31].

Based on the definitions of \mathcal{H}_k and $\widehat{\text{SINR}}_k$, the constraints $\widehat{\text{SINR}}_k(\mathbf{W})/\gamma_k \geq \lambda$ can be rewritten as:

$$F_k(\delta_k, \lambda) \geq 0 \text{ for all } \delta_k \text{ such that } \delta_k^H \delta_k \leq \epsilon_k^2, \quad \forall k, \quad (20)$$

where

$$F_k(\delta_k, \lambda) := (\hat{\mathbf{h}}_k + \delta_k)^H \left(\frac{\mathbf{w}_k \mathbf{w}_k^H}{\lambda \gamma_k} - \sum_{l=1, l \neq k}^K \mathbf{w}_l \mathbf{w}_l^H \right) \times (\hat{\mathbf{h}}_k + \delta_k) - \sigma_k^2.$$

Next, we will show that the infinitely many constraints (20) have an equivalent LMI representation. To this end, we resort to the well-known S-procedure in optimization. This procedure provides an LMI representation for a robust quadratic constraint over an uncertainty set defined by one or two quadratic inequalities. For convenience, we cite the S-procedure in [30] and [29] below:

Lemma 3. (S-procedure) Let \mathbf{A} and \mathbf{B} be two $n \times n$ Hermitian matrices, $\mathbf{c} \in \mathbb{C}^n$ and $d \in \mathbb{R}$. Then the following two conditions are equivalent:

- (i) $\mathbf{x}^H \mathbf{A} \mathbf{x} + \mathbf{c}^H \mathbf{x} + \mathbf{x}^H \mathbf{c} + d \geq 0$ for all $\mathbf{x} \in \mathbb{C}^n$ such that $\mathbf{x}^H \mathbf{B} \mathbf{x} \leq 1$;
- (ii) There exists a $t \in \mathbb{R}$ such that

$$t \geq 0, \quad \begin{bmatrix} \mathbf{A} + t\mathbf{B} & \mathbf{c} \\ \mathbf{c}^H & d - t \end{bmatrix} \succeq \mathbf{0}.$$

By applying Lemma 3, problem (19) can be reformulated as an SDP with rank constraints. Define $\mathbf{X}_k := \mathbf{w}_k \mathbf{w}_k^H$, $\forall k$. Then $\mathbf{X}_k \succeq \mathbf{0}$ and $\text{rank}(\mathbf{X}_k) = 1$. Using the above S-procedure, (20) can be transformed as $\mathbf{\Gamma}_k(\lambda) \succeq \mathbf{0}$ with respect to λ :

$$\mathbf{\Gamma}_k(\lambda) := \begin{bmatrix} \mathbf{Y}_k(\lambda) + t_k \mathbf{I} & \mathbf{Y}_k(\lambda) \hat{\mathbf{h}}_k \\ \hat{\mathbf{h}}_k^H [\mathbf{Y}_k(\lambda)]^H & \hat{\mathbf{h}}_k^H \mathbf{Y}_k(\lambda) \hat{\mathbf{h}}_k - \sigma_k^2 - t_k \epsilon_k^2 \end{bmatrix} \succeq \mathbf{0},$$

where $t_k \geq 0$, $\forall k$, and

$$\mathbf{Y}_k(\lambda) := \frac{1}{\lambda \gamma_k} \mathbf{X}_k - \sum_{l=1, l \neq k}^K \mathbf{X}_l, \quad \forall k. \quad (21)$$

Using the property that $\mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k = \text{tr}(\mathbf{D}_m \mathbf{X}_k)$, we get the following equivalent SDP linear constraints:

$$\sum_{k=1}^K \text{tr}(\mathbf{D}_m \mathbf{X}_k) \leq \alpha P_m, \quad \forall m. \quad (22)$$

Until now, we are ready to show the following proposition:

Proposition 2. (Robust-SP) The problem (19) can be equivalently reformulated as a rank-constrained SDP:

$$\begin{aligned} \tilde{\alpha}^*(\lambda) &= \min_{\{\mathbf{X}_k, t_k, \alpha\}} \alpha \\ \text{s. t. } &\sum_{k=1}^K \text{tr}(\mathbf{D}_m \mathbf{X}_k) \leq \alpha P_m, \quad \forall m, \\ &\mathbf{\Gamma}_k(\lambda) \succeq \mathbf{0}, t_k \geq 0, \quad \forall k, \\ &\mathbf{X}_k \succeq \mathbf{0}, \quad \text{rank}(\mathbf{X}_k) = 1, \quad \forall k. \end{aligned} \quad (23)$$

From Proposition 2, it is evident that, by dropping the rank-one constraints, Robust-SP (23) becomes an SDP which can be solved in a numerically reliable and efficient fashion [29]. Suppose \mathbf{X}_k^* , $\forall k$, are the optimal solutions to the SDP relaxation of (23). If $\mathbf{X}_k^* = \mathbf{w}_k^* \mathbf{w}_k^{*H}$, $\forall k$, i.e., the SDP relaxation is tight, then we obtain the optimal \mathbf{w}_k^* , $\forall k$, for (19). In fact, it was shown in ([15] Theorem 1) that the SDP relaxation under a sum-power constraint always admits a rank-one optimal solution when ϵ_k , $\forall k$, are sufficiently small. Likewise, we extend the approach to establish the following lemma for (23) under per-antenna power constraints:

Lemma 4. Suppose that for some choice of channel uncertainty bounds $\bar{\epsilon} = [\bar{\epsilon}_1, \dots, \bar{\epsilon}_K]^T$, the SDP relaxation of (23) is feasible. Define the set

$$\begin{aligned} \Omega(\bar{\epsilon}) &:= \left\{ \epsilon = [\epsilon_1, \dots, \epsilon_K]^T \mid \epsilon_k \leq \bar{\epsilon}_k, \text{ and } \epsilon_k \right. \\ &\quad \left. < \sqrt{\frac{\lambda \gamma_k \sigma_k^2 \min_m(q_m)}{V(\bar{\epsilon})}}, \forall k \right\}, \end{aligned} \quad (24)$$

where q_m , $\forall m$, are the optimal Lagrange multipliers corresponding to the per-antenna power constraints in (23), and $V(\bar{\epsilon})/(\sum_{m=1}^M P_m)$ is equal to the optimal value for the SDP relaxation of (23) with given $\bar{\epsilon}$. Then, for any vector $\epsilon \in \Omega(\bar{\epsilon})$, we have $\text{rank}(\mathbf{X}_k^*) = 1$, $\forall k$.

Proof. Please see Appendix 3.

On the other hand, for the case of large ϵ_k , $\forall k$, the existence of rank-one optimal solutions for the SDP relaxation of (23) cannot be provably guaranteed. Hence, the exact optimal solution to (19) may not be constructed from \mathbf{X}_k^* with possibly a rank greater than one. In this case, randomization is a widely adopted method to obtain a feasible rank-one approximate matrix solution from the SDP relaxation. Specifically, a Gaussian randomization strategy [31] can be applied to get a vector \mathbf{w}_k^* from \mathbf{X}_k^* , $\forall k$, to nicely approximate the solution to (19). \square

4.3 The Robust-LPM design

As shown in the previous subsection, when the CSI uncertainty bounds ϵ_k , $\forall k$, are small to some extent, the SDP relaxation of Robust-SP (23) is globally optimal. However,

it remains open whether (19) has an equivalent convex reformulation (i.e., there always exists a tight SDP relaxation) in general. As a compromise, we consider another interesting reformulation based on more conservative robust SINR constraints compared to those in (19).

The robust SINR constraints in (19) can be rewritten as:

$$\frac{1}{\sqrt{\lambda}\gamma_k} |\mathbf{h}_k^H \mathbf{w}_k| \geq \sqrt{\sum_{i=1, i \neq k}^K \mathbf{w}_i^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{w}_i + \sigma_k^2}, \quad (25)$$

for $\|\delta_k\| \leq \epsilon_k$, where $\delta_k = \mathbf{h}_k - \hat{\mathbf{h}}_k$, $\forall k$. Consider the following robust SINR constraints which appear a bit conservative:

$$\frac{1}{\sqrt{\lambda}\gamma_k} \mathcal{R}(\mathbf{h}_k^H \mathbf{w}_k) \geq \sqrt{\sum_{i=1, i \neq k}^K \mathbf{w}_i^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{w}_i + \sigma_k^2}, \quad (26)$$

for $\|\delta_k\| \leq \epsilon_k$. Since $\mathcal{R}(\mathbf{h}_k^H \mathbf{w}_k) \leq |\mathbf{h}_k^H \mathbf{w}_k|$, the set of \mathbf{w}_k defined by (26) is always contained in that by (25).

With the conservative SINR constraints (26), we consider the following robust beamforming problem formulation:

$$\begin{aligned} \hat{\alpha}^*(\lambda) &= \min_{\mathbf{W}, \alpha} \alpha \\ \text{s. t. } &\sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq \alpha P_m, \forall m, \\ &\frac{1}{\sqrt{\lambda}\gamma_k} \mathcal{R}(\mathbf{h}_k^H \mathbf{w}_k) \geq \sqrt{\sum_{i=1, i \neq k}^K \mathbf{w}_i^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{w}_i + \sigma_k^2}, \forall k, \\ &\mathbf{h}_k = \hat{\mathbf{h}}_k + \delta_k, \forall \|\delta_k\| \leq \epsilon_k, \forall k. \end{aligned} \quad (27)$$

The problem (27) is a semi-infinite SOCP and hence convex, but it is still not easy to be solved efficiently in the current form. Nevertheless, the conservative SINR constraints (26) can be reformulated as LPMs, where the set of each LPM actually forms a convex cone [32]. Furthermore, the recent elegant result in [33] shows that one can construct an LMI to describe a cone of LPM. Resorting to such an LMI representation, the intended beamforming problem (27) can be reformulated as a convex SDP.

To begin with, we define some notations. The n -dimensional SOC (which is also termed Lorentz cone) is defined as [33]:

$$\mathbb{L}^n = \left\{ \mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2^2 + \dots + x_n^2} \right\}.$$

The real-valued $2M \times 2M$ matrices $\tilde{\mathbf{D}}_m$, $\forall m$, are defined as

$$\tilde{\mathbf{D}}_m := \begin{bmatrix} \mathcal{R}(\mathbf{D}_m) & -\mathcal{I}(\mathbf{D}_m) \\ \mathcal{I}(\mathbf{D}_m) & \mathcal{R}(\mathbf{D}_m) \end{bmatrix} \in \mathbb{R}^{2M \times 2M},$$

and the real-valued $2K \times (2M+1)$ matrices $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$, $\forall k$, are Lorentz positive and satisfy that

$$\mathbf{B}_k(\lambda, \tilde{\mathbf{w}}) \mathbf{y}_k \in \mathbb{L}^{2K}, \forall \mathbf{y}_k \in \mathbb{L}^{2M+1}.$$

Given $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_P]^T$ (with $P, Q \geq 3$, $\mathbf{g}_p \in \mathbb{R}^Q$, $\forall p = 1, \dots, P$), $\hat{\mathbf{A}}(\mathbf{G})$ is generated by the P arrow matrices $\{\mathbf{A}(\mathbf{g}_1), \dots, \mathbf{A}(\mathbf{g}_P)\}$, i.e.,

$$\hat{\mathbf{A}}(\mathbf{G}) := \begin{bmatrix} \mathbf{A}(\mathbf{g}_0) & \mathbf{A}(\mathbf{g}_3) & \mathbf{A}(\mathbf{g}_4) & \cdot & \mathbf{A}(\mathbf{g}_p) \\ \mathbf{A}(\mathbf{g}_3) & \mathbf{A}(\mathbf{g}_{-1}) & \mathbf{0} & \cdot & \mathbf{0} \\ \mathbf{A}(\mathbf{g}_4) & \mathbf{0} & \mathbf{A}(\mathbf{g}_{-1}) & \cdot & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}(\mathbf{g}_p) & \mathbf{0} & \mathbf{0} & \cdot & \mathbf{A}(\mathbf{g}_{-1}) \end{bmatrix}, \quad (28)$$

where $\mathbf{g}_0 = \mathbf{g}_1 + \mathbf{g}_2$, $\mathbf{g}_{-1} = \mathbf{g}_1 - \mathbf{g}_2$, $\mathbf{g}_p = [g_{p1}, g_{p2}, \dots, g_{pQ}]^T$, $\forall p$, and

$$\mathbf{A}(\mathbf{g}_p) := \begin{bmatrix} g_{p1} + g_{p2} & g_{p3} & g_{p4} & \cdot & g_{pQ} \\ g_{p3} & g_{p1} - g_{p2} & 0 & \cdot & 0 \\ g_{p4} & 0 & g_{p1} - g_{p2} & \cdot & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{pQ} & 0 & 0 & \cdot & g_{p1} - g_{p2} \end{bmatrix}.$$

With these notations, we can mimic the approach in [32] to show the following proposition.

Proposition 3. (Robust-LPM) The problem (27) can be reformulated as an SDP:

$$\begin{aligned} \hat{\alpha}^*(\lambda) &= \min_{\{\mathbf{Z}_k\}, \tilde{\mathbf{w}}, \alpha} \alpha \\ \text{s. t. } &\begin{bmatrix} \sqrt{\alpha P_m} \\ (\mathbf{I}_K \otimes \tilde{\mathbf{D}}_m) \tilde{\mathbf{w}} \end{bmatrix} \in \mathbb{L}^{2MK+1}, \forall m, \\ &\hat{\mathbf{A}}(\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})) + \mathbf{Z}_k \geq \mathbf{0}, \forall k, \\ &\mathbf{Z}_k \in \mathcal{L}_{2K-1, 2M}^\perp, \forall k, \end{aligned} \quad (29)$$

where $\tilde{\mathbf{w}} := [\mathbf{w}_{11}^T, \mathbf{w}_{12}^T, \dots, \mathbf{w}_{K1}^T, \mathbf{w}_{K2}^T]^T$, $\mathbf{w}_{k1} = \mathcal{R}(\mathbf{w}_k)$, $\mathbf{w}_{k2} = \mathcal{I}(\mathbf{w}_k)$, $\forall k$, and $\mathcal{L}_{2K-1, 2M}^\perp$ denotes a linear subspace as in (58).

Proof. Please see Appendix 4.

Proposition 3 shows that the robust beamforming problem (27) can be reformulated as an equivalent SDP (29). We highlight that (29) can be efficiently solved using a handy solver (like SeDuMi [29]) with a guaranteed globally optimal solution. This is different in nature from the Robust-SP (23), where the original problem (19) is solved through the SDP relaxation technique (again noting that there is no mathematical proof for the zero gap between (19) and the SDP relaxation of (23) in general). Since more conservative SINR constraints are employed, it is clear that $\hat{\alpha}^*(\lambda) \geq \tilde{\alpha}^*(\lambda)$, i.e., the proposed Robust-LPM might yield a sub-optimal solution to (19). However, simulation results in Section 6 show that Robust-LPM can always provide an (near-)optimal solution.

Before leaving this section, we would like to remark that Robust-LPM (29) includes larger size matrices in the LMI constraints than Robust-SP (19), which means that the Robust-LPM has higher computational complexity than the SDP relaxation of Robust-SP (i.e., removing the

rank-one constraints). However, as a trade-off, the global optimality for Robust-LPM can be achieved while the global optimality for Robust-SP cannot be always guaranteed (i.e., there is a positive gap between the SDP relaxation and the problem itself). Both Robust-LPM and the relaxed Robust-SP can be efficiently solved by interior methods, practically using a software like CVX. According to Boyd and Vandenberghe ([29] Chapter 11.7), the computational complexities of the interior-point method for solving Robust-LPM and the relaxed Robust-SP are $\mathcal{O}(K^7 M^{6.5})$ and $\mathcal{O}(\sqrt{KM} + M^2 K^3 M^6)$, respectively. \square

5 Worst-case weighted sum-rate maximization

In this section, we extend the proposed approach to the optimal robust beamforming for the worst-case WSR maximization under per-antenna power constraints. Relying on the MP method [5–7, 16, 18, 19], we show that an efficient algorithm can be developed for the worst-case WSR maximization through solving a sequence of robust max-min SINR problems.

5.1 Problem formulation

With $\widetilde{\text{SINR}}_k(\mathbf{W}) = \min_{\mathbf{h}_k \in \mathcal{H}_k} \frac{|\mathbf{h}_k^H \mathbf{w}_k|}{\sum_{l \neq k} |\mathbf{h}_k^H \mathbf{w}_l|^2 + \sigma_k^2}$ defined in (15), the achievable worst-case rate of user k can be expressed as:

$$r_k(\mathbf{W}) = \log_2 \left(1 + \widetilde{\text{SINR}}_k(\mathbf{W}) \right) \text{ bits/s/Hz.} \quad (30)$$

Let μ_k denote the priority weight for user k . An optimization problem with the objective function of maximizing the worst-case WSR under the per-antenna power constraints can be formulated as:

$$\begin{aligned} \max_{\mathbf{W}} \quad & \sum_{k=1}^K \mu_k \log_2 \left(1 + \widetilde{\text{SINR}}_k(\mathbf{W}) \right) \\ \text{s. t.} \quad & \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq P_m, \forall m. \end{aligned} \quad (31)$$

The problem (31) is non-convex due to the heavily coupled mutual-interference terms in its objective function. However, it can be next converted to a standard MP problem over a norm set and can be solved efficiently.

Firstly, we introduce some definitions. A set \mathcal{S} is called *normal* if $\mathbf{z}' \leq \mathbf{z}$ and $\mathbf{z} \in \mathcal{S}$ implies $\mathbf{z}' \in \mathcal{S}$; and a set \mathcal{S} is called *reverse normal* if $\mathbf{z}' \geq \mathbf{z}$ and $\mathbf{z} \in \mathcal{S}$ implies $\mathbf{z}' \in \mathcal{S}$. A *box* $[\mathbf{a}, \mathbf{b}]$ is defined as the set of all \mathbf{z} such that $\mathbf{a} \leq \mathbf{z} \leq \mathbf{b}$. A vector $\mathbf{y} \in \mathbb{R}_+^K$ is an upper-boundary point of set \mathcal{S} if $\alpha \mathbf{y} \in \mathcal{S}$, for any scalar $\alpha < 1$, and $\alpha \mathbf{y} \notin \mathcal{S}, \forall \alpha > 1$. The set of upper-boundary points of \mathcal{S} is called *upper-boundary* of \mathcal{S} and it is denoted by $\partial^+ \mathcal{S}$.

Now define the set $\mathcal{W} := \{\mathbf{W} \mid \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k \leq P_m, \forall m\}$. Introducing an auxiliary vector $\mathbf{z} = [z_1, \dots, z_K]^T$, we can equivalently reformulate (31) as:

$$\max_{\mathbf{z} \in \mathcal{Z}} \Phi(\mathbf{z}) := \sum_{k=1}^K \mu_k \log_2(z_k), \quad (32)$$

where the feasible set is

$$\mathcal{Z} := \left\{ \mathbf{z} \mid 1 \leq z_k \leq 1 + \widetilde{\text{SINR}}_k(\mathbf{W}), \forall k, \forall \mathbf{W} \in \mathcal{W} \right\}. \quad (33)$$

Since $\Phi(\mathbf{z})$ is a strictly increasing function in each entry of \mathbf{z} , the optimal solution \mathbf{z}^* for (32) must satisfy: $z_k^* = 1 + \widetilde{\text{SINR}}_k(\mathbf{W}^*)$, $\forall k$, for a certain $\mathbf{W}^* \in \mathcal{W}$; and such \mathbf{W}^* is clearly the optimal solution to the original problem (31).

Further define the following two sets

$$\begin{aligned} \mathcal{G} &:= \left\{ \mathbf{z} \mid 0 \leq z_k \leq 1 + \widetilde{\text{SINR}}_k(\mathbf{W}), \forall k, \forall \mathbf{W} \in \mathcal{W} \right\}, \\ \mathcal{H} &:= \{ \mathbf{z} \mid z_k \geq 1, \forall k \}, \end{aligned}$$

and let $\mathbf{a}(\mathbf{W}) := [1 + \widetilde{\text{SINR}}_1(\mathbf{W}), \dots, 1 + \widetilde{\text{SINR}}_K(\mathbf{W})]^T$ for any $\mathbf{W} \in \mathcal{W}$. Then $\mathcal{G} = \cup_{\mathbf{W} \in \mathcal{W}} [\mathbf{0}, \mathbf{a}(\mathbf{W})]$ is the union of infinitely many normal boxes; clearly, \mathcal{G} is normal [34]. Let $\mathbf{b} := [b_1, \dots, b_K]^T$, where

$$b_k := 1 + \frac{\sum_{m=1}^M P_m \left(\|\hat{\mathbf{h}}_k\|^2 + \|\epsilon_k\|^2 \right)}{\sigma_k^2}, \forall k. \quad (34)$$

It clearly holds: $1 + \widetilde{\text{SINR}}_k(\mathbf{W}) \leq b_k, \forall k$. Therefore, $\mathcal{G} \subset [\mathbf{0}, \mathbf{b}]$ is a compact normal set with non-empty interior. It is also clear that \mathcal{H} is a reverse normal set. It then follows that (32) is a standard MP [34]:

$$\max_{\mathbf{z}} \Phi(\mathbf{z}), \quad \text{s. t. } \mathbf{z} \in (\mathcal{G} \cap \mathcal{H}), \quad (35)$$

where we maximize an increasing function $\Phi(\mathbf{z})$ over the intersection of a compact normal set \mathcal{G} and a reverse normal set \mathcal{H} . For the MP, the maximum is attained on $\partial^+(\mathcal{G} \cap \mathcal{H})$.

5.2 The POA algorithm

Based on the separation property³ of normal sets [34], a *polyblock outer approximation* (POA) algorithm can be employed to efficiently find the *globally* optimal solution for an MP. For any finite vector set $\mathcal{T} := \{\mathbf{v}_i, i = 1, \dots, I\}$, the union of all the boxes $[\mathbf{0}, \mathbf{v}_i], \forall i$, is a *polyblock*. The basic idea is to approximate the feasible set $\mathcal{G} \cap \mathcal{H}$ by a polyblock. Specifically, we construct a nested sequence of polyblocks: $\mathcal{P}_1 \supset \mathcal{P}_2 \supset \dots \supset \mathcal{P}_n \supset \dots \supset \mathcal{G} \cap \mathcal{H}$ in such a way that $\max_{\mathbf{z} \in \mathcal{P}_n} \Phi(\mathbf{z}) \searrow \max_{\mathbf{z} \in (\mathcal{G} \cap \mathcal{H})} \Phi(\mathbf{z})$, where \mathcal{P}_n denotes the polyblock generated at the n th iteration and “ \searrow ” denotes convergence from above.

A vertex $\mathbf{v}_i \in \mathcal{T}$ is called *proper* if there does not exist another $\mathbf{v}_j \in \mathcal{T}$ such that $\mathbf{v}_j \geq \mathbf{v}_i$. The maximum of an increasing function over a polyblock is attained at one of its proper vertices. Hence, at the n th iteration, the maximizer for $\Phi(\mathbf{z})$ over \mathcal{P}_n is obtained as:

$$\mathbf{z}^n := \arg \max_{\mathbf{z} \in \mathcal{T}_n} \Phi(\mathbf{z}), \quad (36)$$

where \mathcal{T}_n is the (finite) proper vertex set of \mathcal{P}_n . Note that \mathbf{z}^n can be simply obtained by searching over the finite entries of \mathcal{T}_n . If $\mathbf{z}^n \in (\mathcal{G} \cap \mathcal{H})$, then it solves (35).

Otherwise, we construct the next polyblock \mathcal{P}_{n+1} contained in $\mathcal{P}_n \setminus \{\mathbf{z}^n\}$ but still enclosing $\mathcal{G} \cap \mathcal{H}$, and continue the process.

To construct new polyblock \mathcal{P}_{n+1} from \mathcal{P}_n , it requires computing the *projection* point $\pi_{\mathcal{G}}(\mathbf{z}^n)$, which is defined as the unique point where the halfline from $\mathbf{0}$ through \mathbf{z} meets $\partial^+ \mathcal{G}$, i.e., $\pi_{\mathcal{G}}(\mathbf{z}) = \lambda \mathbf{z}$ and $\lambda = \max\{\alpha \mid \alpha \mathbf{z} \in \mathcal{G}\}$. Let $\mathbf{y}^n = [y_1^n, \dots, y_K^n]^T := \pi_{\mathcal{G}}(\mathbf{z}^n)$, and

$$\mathbf{z}^{n,k} := \mathbf{z}^n - (z_k^n - y_k^n) \mathbf{e}^k, \forall k, \quad (37)$$

where \mathbf{e}^k is a unit vector with the only non-zero number (i.e., “1”) in the k th entry. Hence, $\mathbf{z}^{n,k}$ can be obtained through replacing z_k^n by y_k^n , and $\mathbf{y}^n \leq \mathbf{z}^{n,k} \leq \mathbf{z}^n, \forall k$.

Let \mathcal{T}_{n+1} be the set obtained from \mathcal{T}_n by replacing the vertex \mathbf{z}^n with K new vertices $\mathbf{z}^{n,k}$ and then removing the improper vertices; i.e., $\mathcal{T}_{n+1} = (\mathcal{T}_n \setminus \{\mathbf{z}^n\}) \cup \{\mathbf{z}^{n,k} \mid \mathbf{z}^{n,k} \text{ is proper}\}$. Since $\mathbf{z}^* \in \mathcal{H}$, we can further reduce the vertex set $\mathcal{T}_{n+1} = \mathcal{T}_{n+1} \cap \mathcal{H}$. With such a vertex set \mathcal{T}_{n+1} , we have ([34] Proposition 17):

Lemma 5. *The polyblock \mathcal{P}_{n+1} with vertex set \mathcal{T}_{n+1} satisfies $(\mathcal{G} \cap \mathcal{H}) \subset \mathcal{P}_{n+1} \subset \mathcal{P}_n$.*

Lemma 5 provides the direction for construction of \mathcal{P}_{n+1} from \mathcal{P}_n to approximate (33) from outside with increasing accuracy. To this end, we need to find $\mathbf{y}^n = \pi_{\mathcal{G}}(\mathbf{z}^n) = \lambda^n \mathbf{z}^n$, which is determined by solving:

$$\begin{aligned} \lambda^n &= \max \{ \alpha \mid \alpha \mathbf{z}^n \in \mathcal{G} \} \\ &= \max \left\{ \alpha \mid \alpha \leq \min_{1 \leq k \leq K} \frac{1 + \widetilde{\text{SINR}}_k(\mathbf{W})}{z_k^n}, \forall \mathbf{W} \in \mathcal{W} \right\} \\ &= \max_{\mathbf{W} \in \mathcal{W}} \min_{1 \leq k \leq K} \frac{1 + \widetilde{\text{SINR}}_k(\mathbf{W})}{z_k^n}. \end{aligned}$$

This leads to the following max-min SINR balancing problem:

$$\begin{aligned} \lambda^n &= \max_{\mathbf{W}} \min_{1 \leq k \leq K} \frac{1 + \widetilde{\text{SINR}}_k(\mathbf{W})}{z_k^n} \\ \text{s. t. } \sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k &\leq P_m, \forall m, \end{aligned} \quad (38)$$

and its “dual” min-max power problem with given λ^n is

$$\begin{aligned} \tilde{\alpha}^*(\lambda) &= \min_{\mathbf{W}} \max_{1 \leq m \leq M} \frac{\sum_{k=1}^K \mathbf{w}_k^H \mathbf{D}_m \mathbf{w}_k}{P_m} \\ \text{s. t. } \frac{1 + \widetilde{\text{SINR}}_k(\mathbf{W})}{z_k^n} &\geq \lambda^n, \forall k. \end{aligned} \quad (39)$$

Problem (39) can be solved either by Robust-SP in Section 4.2 or by Robust-LPM in Section 4.3. The proposed Algorithm 1 can be applied to find the optimal value λ^n for (38).

The proposed POA algorithm to find an ξ -optimal solution for (31) is summarized in Algorithm 2. A key requirement for provable convergence is that \mathbf{z} is lower bounded by a strictly positive vector [34]. Since $\mathbf{z} \geq \mathbf{1} > \mathbf{0}$ in (32), it readily follows from ([34] Theorem 1) that:

Proposition 4. *Algorithm 2 globally converges to an ξ -optimal solution for (31) and (32).*

Algorithm 2 for the worst-case WSR Maximization (31)

- 1: **Initialize:** Select a tolerance level $\xi > 0$;
 - 2: Let $n=0$, $\mathcal{T}_0 = \{\mathbf{b}\}$, and CBV (current best value) = 0.
 - 3: **Repeat:** Let $\mathbf{z}^n = \arg \max_{\mathbf{z} \in \mathcal{T}_n} \Phi(\mathbf{z})$;
 - 4: For \mathbf{z}^n , use Algorithm 1 to solve (38) to obtain λ^n , the corresponding \mathbf{W}^* , as well as $\mathbf{y}^n = \lambda^n \mathbf{z}^n$.
 - 5: If $\mathbf{y}^n \in \mathcal{H}$ and $\Phi(\mathbf{y}^n) > \text{CBV}$, then $\text{CBV} = \Phi(\mathbf{y}^n)$, $\bar{\mathbf{z}} = \mathbf{y}^n$, and $\bar{\mathbf{W}} = \mathbf{W}^*$.
 - 6: Let $\mathbf{z}^{n,k} = \mathbf{z}^n - (z_k^n - y_k^n) \mathbf{e}^k, \forall k$, and $\mathcal{T}_{n+1} = [(\mathcal{T}_n \setminus \{\mathbf{z}^n\}) \cup \{\mathbf{z}^{n,k} \mid \mathbf{z}^{n,k} \text{ is proper}\}] \cap \mathcal{H}$.
 - 7: Further remove from \mathcal{T}_{n+1} any $\mathbf{v}_i \in \mathcal{T}_{n+1}$ satisfying $\Phi(\mathbf{v}_i) \leq \text{CBV}(1 + \xi)$.
 - 8: Set $n = n + 1$.
 - 9: **Until** $\mathcal{T}_n = \emptyset$.
 - 10: **Output:** $\bar{\mathbf{z}}$ as the ξ -optimal solution for (32) and $\bar{\mathbf{W}}$ the solution for (31).
-

Proposition 4 establishes that Algorithm 2 can yield the optimal beamformers for (31) with guaranteed convergence and global ξ -optimality. Note that the worst-case WSR maximization (31) is in fact NP-hard. The proposed POA method (i.e., Algorithm 2) is essentially a smart branch-and-bound approach, which does not have a worst-case polynomial-time complexity [34]. However, extensive numerical examples have shown that this type of algorithm can solve general MP problems of dimensions 10–15 (while the problems of such dimensions are already very hard to solve by the standard approximation tools) [5, 6, 35]. For those small/medium size problems, the complexity with the algorithm may be affordable for practical implementation. On the other hand, for problems with even larger dimensions, the algorithm may be only suitable for benchmarking purposes.

It is worth noting that similar POA algorithms were adopted to find the optimal beamforming designs for the WSR maximization in other different setups [6, 7, 18, 19, 35]. The proposed Algorithm 2 can be also applied to the optimal robust beamforming designs under other general criteria where the objective functions, e.g., the sum of mean square errors or sum of bit error rates, are monotonic functions of the SINRs; see a unified framework [18, 35].

6 Numerical results

6.1 Simulation setting

Consider a MISO downlink system where the BS, equipped with M antennas, serves K single-antenna users. Each transmit antenna has an equal power budget, i.e.,

$P_m = P/M, \forall m$, where P is the total power budget at the BS. Assume a normalized channel estimation $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. The CSI uncertainty bound $\epsilon_k = \epsilon \|\hat{\mathbf{h}}_k\|, \forall k$, for a given ϵ in (14). We assume that $n_k \sim \mathcal{CN}(0, \sigma^2), \forall k$, and $\sigma^2 = 0.1$. Define $\text{SNR} = P/\sigma^2$, and each user has an equal SINR target, i.e., $\gamma_k = \text{SNR}/K, \forall k$. The tolerance levels are set as $\delta = 0.01$ and $\xi = 0.01$ in Algorithms 1 and 2, respectively.

6.2 Properties of $\alpha(\lambda)$, $\hat{\alpha}(\lambda)$, and $\tilde{\alpha}(\lambda)$

For one given channel realization with $M = 4$ and $K = 3$, Fig. 2 shows the curves of the per-antenna min-max power functions $\alpha^*(\lambda)$ in (10) with perfect CSI, $\tilde{\alpha}^*(\lambda)$ in (23) of Robust-SP, and $\hat{\alpha}^*(\lambda)$ in (29) of Robust-LPM with respect to the SINR-to-target ratio requirement λ . The input SNR is chosen as $\text{SNR} = 10$ dB, and the channel uncertainty bound norm is $\epsilon = 0.2$. The monotonicity of $\alpha^*(\lambda)$, $\tilde{\alpha}^*(\lambda)$, and $\hat{\alpha}^*(\lambda)$ is clearly shown in Fig. 2. Robust-SP and Robust-LPM achieve almost the same value, i.e., $\tilde{\alpha}^*(\lambda) \approx \hat{\alpha}^*(\lambda)$. It shows that the values of $\alpha^*(\lambda)$ are smaller than those of the robust ones which consider the channel uncertainty. It means that the robust beamforming design requires more power than the beamforming design with perfect CSI under the same SINR constraints to account for the CSI uncertainty. Besides, as discussed in Sections 3 and 4, the optimal beamforming designs can be obtained by solving $\alpha^*(\lambda) = 1$, $\tilde{\alpha}^*(\lambda) = 1$, or $\hat{\alpha}^*(\lambda) = 1$.

6.3 Average minimum SINR

To gauge the average minimum SINR performance of different designs with uncertainty CSI, we realize 100 independent simulation runs. For each simulation run, we first generate $\hat{\mathbf{h}}_k, \forall k$, randomly. Based on (13) in Section 4.1,

we obtain 1000 true channel realizations $\mathbf{h}_k = \hat{\mathbf{h}}_k + \delta_k$ (where $\|\delta_k\| \leq \epsilon \|\hat{\mathbf{h}}_k\|$), and compute the average minimum SINR over the channel realizations for each input SNR. Then, the SINR results are averaged over the 100 simulation runs. Figures 3 and 4 demonstrate the average minimum SINR performance among $K = 3$ users for $M = 4$ and $M = 6$, respectively, where $\epsilon = 0.2$. It is observed that the proposed Algorithm 1 based on Robust-LPM achieves almost the same average minimum SINR performance as that based on Robust-SP.

We also evaluate the rank-relaxation method in [16] under per-antenna power constraints, where the robust beamformers are obtained using a bisection search with the aid of SDP and rank-relaxation, and all the iterations that return rank-one solutions are stored. In Figs. 3 and 4, we observe that Robust-SP/Robust-LPM significantly outperforms the method of [16] in high SNR regimes. For example, at $\text{SNR} = 10$ dB in Fig. 3, the proposed designs achieve about 4 dB greater SINR gains than the method of [16]. It implies, in high SNR regimes, the tightness of rank relaxation becomes invalid, and thus a randomization is necessary. This is also shown in Lemma 4, when SNR (which is proportional to $V(\bar{\epsilon})$) increases, the CSI uncertainty set that guarantees the tightness becomes smaller.

In addition, the performances of NonRobust design, the minimum mean square error (MMSE) design in [4], and the Moore-Penrose ZF design in [23] are also included in Figs. 3 and 4. For these three designs, although the true channel is \mathbf{h}_k , the calculation of beamformers is processed by regarding the estimated channel $\hat{\mathbf{h}}_k$ as the true channel. Note that in order to implement the Moore-Penrose ZF beamforming design, we assume that $M \geq K$. The optimal

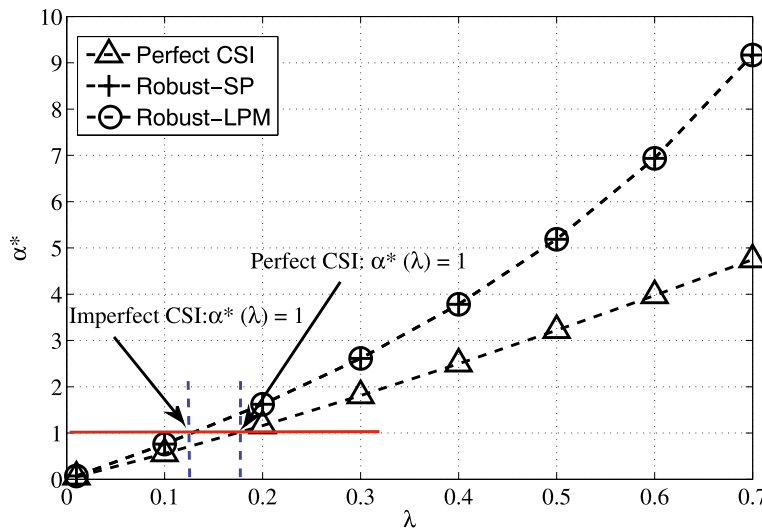


Fig. 2 Illustration of $\alpha^*(\lambda)$ in (8), $\tilde{\alpha}^*(\lambda)$ in (23), and $\hat{\alpha}^*(\lambda)$ in (29), $M = 4$, $K = 3$, $\text{SNR} = 10$ dB

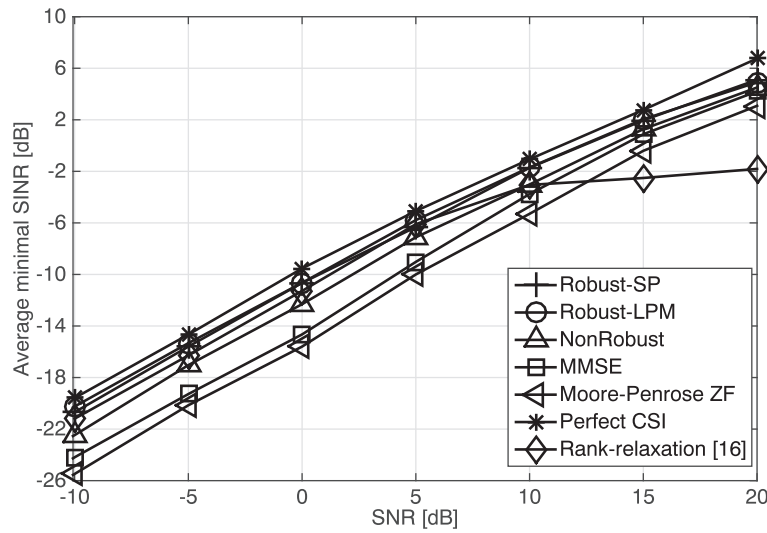


Fig. 3 Average minimum SINR vs SNR with $\epsilon = 0.2$, $M = 4$, and $K = 3$

performance under perfect CSI is also provided as an upper bound for all designs. Figures 3 and 4 show that the proposed Robust-SP and Robust-LPM designs outperform NonRobust in the whole SNR regimes. This is expected because NonRobust regards the estimated channel $\hat{\mathbf{h}}_k$ as the true CSI \mathbf{h}_k in calculating the beamformers. If the perfect CSI were available, the more “aggressive” NonRobust would have yielded a higher average minimum user SINR than the robust ones (since the robust designs are conservative in the sense that they aim to provide the worse-case guarantees). However, in the presence of the CSI uncertainty, the performance of NonRobust degrades dramatically at most of times since that pretty small

perturbations of uncertainty data can make the nominal optimal solution heavily infeasible and thus meaningless.

The proposed robust designs clearly achieve a larger gain in the average minimum SINR performance than the MMSE and Moore-Penrose ZF designs as shown in Figs. 3 and 4. For example, at SNR = 0 dB, the Robust-SP and the Robust-LPM designs with $M = 4$ achieve about a 4 dB gain in the average minimum SINR performance when compared to the MMSE design. Lastly, we observe that the average minimum SINR performance improves when the number of transmit antennas grows. This is because more antennas lead to more accurate beamformers matched with channel characteristics.

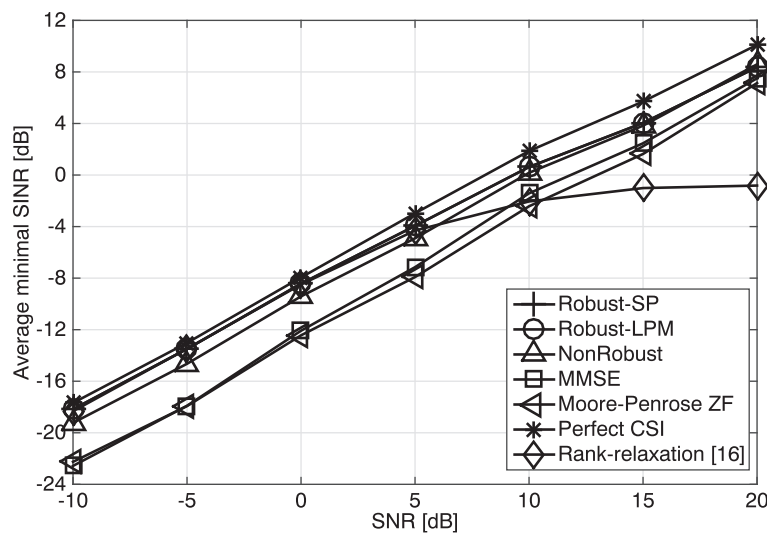


Fig. 4 Average minimum SINR vs SNR with $\epsilon = 0.2$, $M = 6$, and $K = 3$

Next, we evaluate the average minimum SINR performance of various designs under different CSI uncertainty norm bounds with $\text{SNR} = 5$ dB and $K = 3$. Figure 5 demonstrates the average minimum SINR performance for $M = 4$ and $M = 6$. Again, it is seen that the proposed Robust-LPM achieves almost the same performance as the proposed Robust-SP. As expected, the average minimum SINR decreases as the channel uncertainty norm bound ϵ increases. For example, the average minimum SINR of the proposed two designs when $\epsilon = 0.5$ is about 1 dB smaller than those when $\epsilon = 0.1$.

6.4 Average WSR maximization

Figure 6 shows the average WSR values of different designs with $\epsilon = 0.2$, $M = 4$, and $K = 3$. The weights are $\mu_1 = 0.2$, $\mu_2 = 0.3$, and $\mu_3 = 0.5$. The proposed worst-case WSR maximization beamformers are obtained via Algorithm 2 with $\xi = 0.01$. “Robust-SP” and “Robust-LPM” in Fig. 6 denote the max-min SINR solutions in Algorithm 2 are obtained through Algorithm 1 based on Robust-SP and Robust-LPM, respectively. It is observed that the WSR maximization design based on Robust-SP slightly outperforms the one based on Robust-LPM, where the corresponding WSR gap is smaller than 0.1 bits/s/Hz in the SNR regimes of consideration. For comparison, we also demonstrate the performances of the WSR maximization beamforming designs for which the max-min SINR solutions are respectively obtained based on the NonRobust design, the MMSE design [4] and the Moore-Penrose ZF design [23]. As shown in Fig. 6, the proposed robust WSR maximization designs achieve significantly higher average WSR, compared to the NonRobust,

MMSE, and Moore-Penrose ZF ones in the medium and high SNR regimes. For example, at an SNR of 10 dB, the proposed Robust-SP achieves about 0.2 bits/s/Hz and 0.6 bits/s/Hz more average WSR over the NonRobust design and the Moore-Penrose ZF design, respectively.

7 Conclusions

Under the per-antenna power constraints, we proposed an efficient approach to find the robust max-min SINR beamforming designs by solving a sequence of min-max power problems. We developed Robust-SP and Robust-LPM for the robust min-max power problems. Building on the min-max power solutions, a bisection search algorithm was developed to obtain the robust max-min SINR beamformers. Using the max-min SINR solution as a cornerstone, we further proposed an MP method to find the robust beamformers for the worst-case WSR maximization with guaranteed convergence and global optimality. Numerical results demonstrated that the proposed robust designs provide substantial performance improvement over the existing alternatives.

8 Appendices

8.1 Appendix 1: Proof of Lemma 1

Let $\tilde{\mathbf{W}} = [\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_M]$ be the optimal solution for (17) with $\lambda > 0$, such that $\tilde{\alpha}^*(\lambda) = \max_{1 \leq m \leq M} \frac{[\sum_{k=1}^K \tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_k^H]_{m,m}}{P_m}$ and $\frac{\widetilde{\text{SINR}}_k(\tilde{\mathbf{W}})}{\gamma_k} \geq \lambda$.

For another $\lambda' \in [0, \lambda]$, let $\beta = \lambda'/\lambda$. It is clear: $0 < \beta < 1$. we can show that $\sqrt{\beta} \tilde{\mathbf{W}}$ is in the feasible set of (17) with λ' . This is because: $\forall k$,

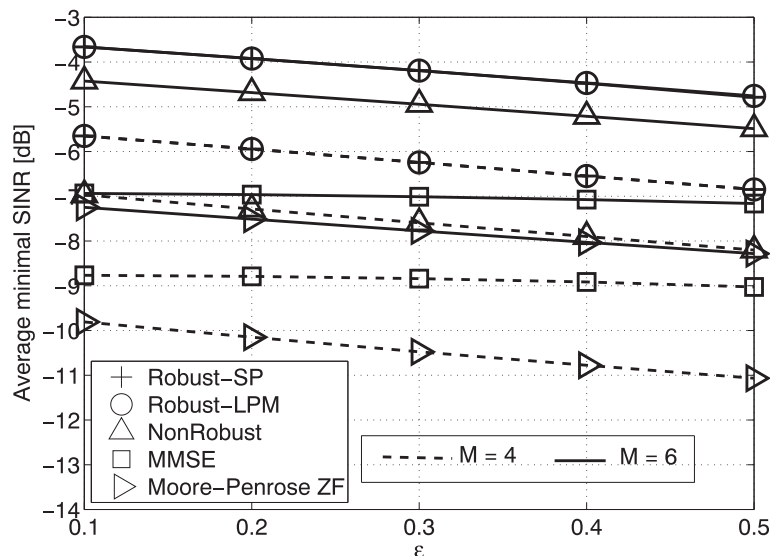


Fig. 5 Average minimum SINR vs ϵ with $\text{SNR} = 5$ dB and $K = 3$

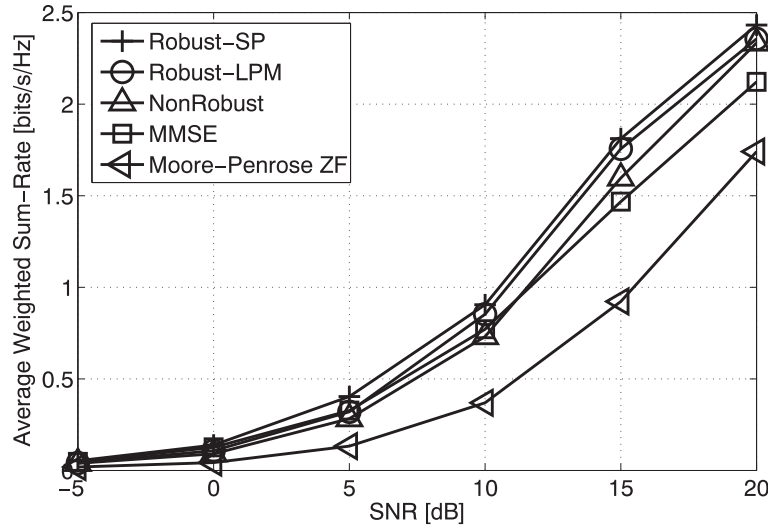


Fig. 6 Average WSR of different beamforming designs with $\epsilon = 0.2$, $M = 4$, and $K = 3$. $\mu_1 = 0.2$, $\mu_2 = 0.3$, and $\mu_3 = 0.5$

$$\begin{aligned} \frac{\widetilde{\text{SINR}}_k(\sqrt{\beta}\tilde{\mathbf{w}})}{\gamma_k} &= \min_{\mathbf{h} \in \mathcal{H}_k} \frac{\beta |\mathbf{h}_k^H \tilde{\mathbf{w}}|^2}{\sum_{l \neq k, l=1}^K (\beta |\mathbf{h}_k^H \tilde{\mathbf{w}}_l|^2) + \sigma_k^2} \\ &\geq \min_{\mathbf{h}_k \in \mathcal{H}_k} \frac{\beta |\mathbf{h}_k^H \tilde{\mathbf{w}}|^2}{\sum_{l \neq k, l=1}^K (|\mathbf{h}_k^H \tilde{\mathbf{w}}_l|^2) + \sigma_k^2} \geq \beta \lambda = \lambda'. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \max_{1 \leq m \leq M} \frac{\left[\sum_{k=1}^K \left(\sqrt{\beta} \tilde{\mathbf{w}}_k \sqrt{\beta} \tilde{\mathbf{w}}_k^H \right) \right]_{m,m}}{P_m} \\ = \beta \max_{1 \leq m \leq M} \frac{\left[\sum_{k=1}^K \left(\tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_k^H \right) \right]_{m,m}}{P_m} = \beta \tilde{\alpha}^*(\lambda). \end{aligned}$$

Therefore, we must have $\tilde{\alpha}^*(\lambda') \leq \beta \tilde{\alpha}^*(\lambda)$. It is easy to see that $\tilde{\alpha}^*(\lambda) > 0$ for any $\lambda > 0$. It in turn implies that $\tilde{\alpha}^*(\lambda') \leq \beta \tilde{\alpha}^*(\lambda) < \tilde{\alpha}^*(\lambda)$ for $0 < \lambda' < \lambda$.

8.2 Appendix 2: Proof of Lemma 2

Since for all $m = 1, \dots, M$,

$$\begin{aligned} \frac{\left[\sum_{k=1}^K \tilde{\mathbf{w}}_k^*(\check{\lambda}) \left[\tilde{\mathbf{w}}_k^*(\check{\lambda}) \right]^H \right]_{m,m}}{P_m} \\ \leq \max_{1 \leq m \leq M} \frac{\left[\sum_{k=1}^K \tilde{\mathbf{w}}_k^*(\check{\lambda}) \left[\tilde{\mathbf{w}}_k^*(\check{\lambda}) \right]^H \right]_{m,m}}{P_m} = \tilde{\alpha}^*(\check{\lambda}) = 1, \end{aligned} \quad (40)$$

then the beamforming matrix $\mathbf{W}^*(\check{\lambda})$ is in the feasible set of (16). This implies:

$$\lambda^* \geq \min_{1 \leq k \leq K} \frac{\widetilde{\text{SINR}}_k(\mathbf{W}^*(\check{\lambda}))}{\gamma_k} \geq \check{\lambda}. \quad (41)$$

Let $\mathbf{W}^* = [\mathbf{w}_1^*, \dots, \mathbf{w}_K^*]$ be the optimal solution of (16). We can show that \mathbf{W}^* is in the feasible set of (17) with λ^* as the minimum $\frac{\widetilde{\text{SINR}}_k(\mathbf{W})}{\gamma_k}$ requirement. This is because $\lambda^* = \min_{1 \leq k \leq K} \frac{\widetilde{\text{SINR}}_k(\mathbf{W}^*)}{\gamma_k}$, or equivalently, $\frac{\widetilde{\text{SINR}}_k(\mathbf{W}^*)}{\gamma_k} \geq \lambda^*, \forall k$. On the other hand, we have $\left[\sum_{k=1}^K \mathbf{w}_k^* \left[\mathbf{w}_k^* \right]^H \right]_{m,m} \leq P_m, \forall m$. Therefore, we have

$$\tilde{\alpha}^*(\lambda^*) \leq \max_{1 \leq m \leq M} \frac{\left[\sum_{k=1}^K \mathbf{w}_k^* \left[\mathbf{w}_k^* \right]^H \right]_{m,m}}{P_m} \leq 1. \quad (42)$$

By Lemma 1, $\tilde{\alpha}^*(\lambda)$ is a strictly increasing function of $\lambda > 0$. The inequality $\lambda^* \geq \check{\lambda}$ in (41) implies:

$$\tilde{\alpha}^*(\lambda^*) \geq \tilde{\alpha}^*(\check{\lambda}) = 1. \quad (43)$$

We have both (42) and (43) only when all the inequalities are satisfied with equality, i.e., $\lambda^* = \check{\lambda}$ and $\mathbf{W}^* = \mathbf{W}^*(\check{\lambda})$.

8.3 Appendix 3: Proof of Lemma 4

We first derive the dual of the downlink beamforming problem (23) without rank-one constraints. For convenience of analysis, we scale the objective function of (23) with $\sum_{m=1}^M P_m$, so that the objective of minimization is the total transmission power $\alpha \sum_{k=1}^K P_m$. Therefore, the problem of interest is expressed as:

$$\begin{aligned} \min_{\{X_k, t_k, \alpha\}} \quad & \alpha \sum_{m=1}^M P_m \\ \text{s. t.} \quad & \sum_{k=1}^K \text{tr}(\mathbf{D}_m \mathbf{X}_k) \leq \alpha P_m, \forall m, \\ & \mathbf{\Gamma}_k(\lambda) \geq \mathbf{0}, t_k \geq 0, \forall k, \\ & \mathbf{X}_k \geq \mathbf{0}, \forall k. \end{aligned} \quad (P_\epsilon)$$

It is evident that problem (P_ϵ) and the SDP relaxation (23) have the same feasible region and the same optimal

solution. We introduce the following dual variables for the corresponding constraints as shown in Table 1.

Let $\mathbf{Q} = \text{diag}(q_1, \dots, q_M)$ and $\Phi = \text{diag}(P_1, \dots, P_M)$, then the dual of this SDP is given by

$$\begin{aligned} & \max_{\{\mathbf{Q}, \mathbf{G}_k, \mathbf{d}_k, \beta_k\}} \sum_{k=1}^K \beta_k \sigma_k^2 \\ & \text{s. t. } \text{tr}(\Phi) \geq \text{tr}(\mathbf{Q}\Phi), \mathbf{Q} \succeq \mathbf{0}, \\ & \mathbf{U}_k := \begin{bmatrix} \mathbf{G}_k & \mathbf{d}_k \\ \mathbf{d}_k^H & \beta_k \end{bmatrix} \succeq \mathbf{0}, \forall k, \\ & \mathbf{Z}_k := \mathbf{Q} - \frac{1}{\lambda \gamma_k} \left(\mathbf{G}_k + \mathbf{d}_k \hat{\mathbf{h}}_k^H + \hat{\mathbf{h}}_k \mathbf{d}_k^H + \beta_k \hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H \right) \\ & + \sum_{l \neq k} \left(\mathbf{G}_l + \mathbf{d}_l \hat{\mathbf{h}}_l^H + \hat{\mathbf{h}}_l \mathbf{d}_l^H + \beta_l \hat{\mathbf{h}}_l \hat{\mathbf{h}}_l^H \right) \succeq \mathbf{0}, \forall k, \\ & c_k := \beta_k \epsilon_k^2 - \text{Tr}(\mathbf{G}_k) \geq 0, \forall k. \end{aligned} \quad (D_\epsilon)$$

Note that the CSI uncertainty bounds $\epsilon = [\epsilon_1, \dots, \epsilon_K]^T$ and λ are regarded as the given parameters for both the primal problem (P_ϵ) and its dual (D_ϵ). Suppose $\{X_k, t_k, \alpha\}$ and $\{\mathbf{Q}, \mathbf{G}_k, \mathbf{d}_k, \beta_k\}$ are the optimal solutions of (P_ϵ) and (D_ϵ), respectively. Similar to the discussions in [15, 20], we have:

- Problem (D_ϵ) is always strictly feasible;
- Suppose the primal problem (P_ϵ) is feasible, then strong duality holds true for the prime (P_ϵ) and its dual (D_ϵ);
- $\beta_k > 0, \forall k$;
- $\mathbf{G}_k + \mathbf{d}_k \hat{\mathbf{h}}_k^H + \hat{\mathbf{h}}_k \mathbf{d}_k^H + \beta_k \hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H \succeq \mathbf{0}, \forall k$;
- $\text{rank}(X_k) \geq 1, \forall k$.

For any $\epsilon = [\epsilon_1, \dots, \epsilon_K]^T \in \Omega(\bar{\epsilon})$, there is $\epsilon_k \leq \bar{\epsilon}_k, \forall k$. This follows that any feasible point of ($P_{\bar{\epsilon}}$) must be a feasible point of (P_ϵ). Hence, (P_ϵ) is solvable. Let $V(\epsilon)$ denote the optimal value of (P_ϵ) (or its dual (D_ϵ)). It is clear that

$$V(\epsilon) \leq V(\bar{\epsilon}). \quad (44)$$

Now, to establish $\text{rank}(X_k) = 1$, we use the complementary slackness condition $\text{Tr}(X_k \mathbf{Z}_k) = 0$ and prove $\text{rank}(\mathbf{Z}_k) = M - 1$. According to the aforementioned properties, we have $\beta_k > 0$ and

$$\begin{aligned} & \mathbf{G}_k + \mathbf{d}_k \hat{\mathbf{h}}_k^H + \hat{\mathbf{h}}_k \mathbf{d}_k^H + \beta_k \hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H \\ & = \left(\mathbf{G}_k - \frac{1}{\beta_k} \mathbf{d}_k \mathbf{d}_k^H \right) + \left(\frac{1}{\sqrt{\beta_k}} \mathbf{d}_k + \sqrt{\beta_k} \hat{\mathbf{h}}_k \right) \\ & \quad \times \left(\frac{1}{\sqrt{\beta_k}} \mathbf{d}_k + \sqrt{\beta_k} \hat{\mathbf{h}}_k \right)^H \succeq \mathbf{0}, \forall k. \end{aligned} \quad (45)$$

Table 1 Dual variables and their corresponding constraints

Dual variable	Constraint
q_m	$\sum_{k=1}^K \text{tr}(\mathbf{D}_m \mathbf{X}_k) - \alpha P_m \leq 0$
\mathbf{U}_k	$\Gamma_k(\lambda) \succeq \mathbf{0}$
c_k	$t_k \geq 0$
\mathbf{Z}_k	$\mathbf{X}_k \succeq \mathbf{0}$

Additionally,

$$\begin{aligned} \text{Tr}(\mathbf{G}_k) & \leq \beta_k \epsilon_k^2 \leq \frac{V(\epsilon)}{\sigma_k^2} \epsilon_k^2 \leq \frac{V(\bar{\epsilon})}{\sigma_k^2} \epsilon_k^2 \\ & < \frac{V(\bar{\epsilon})}{\sigma_k^2} \frac{\lambda \gamma_k \min_m q_m}{V(\bar{\epsilon})} = \lambda \gamma_k \min_m q_m. \end{aligned} \quad (46)$$

This implies that $\text{tr}(\frac{1}{\lambda \gamma_k} \mathbf{G}_k) < \min_m q_m$, which leads to

$$\mathbf{Q} - \frac{1}{\lambda \gamma_k} \mathbf{G}_k \succ \mathbf{0}. \quad (47)$$

Therefore, the $\text{rank}(\mathbf{Z}_k)$ is calculated as:

$$\begin{aligned} \text{rank}(\mathbf{Z}_k) & = \text{rank} \left(\left(\mathbf{Q} - \frac{1}{\lambda \gamma_k} \mathbf{G}_k \right) + \frac{1}{\lambda \gamma_k \beta_k} \mathbf{d}_k \mathbf{d}_k^H \right. \\ & \quad \left. + \sum_{l \neq k} \left(\mathbf{G}_l + \mathbf{d}_l \hat{\mathbf{h}}_l^H + \hat{\mathbf{h}}_l \mathbf{d}_l^H + \beta_l \hat{\mathbf{h}}_l \hat{\mathbf{h}}_l^H \right) \right. \\ & \quad \left. - \left(\frac{1}{\sqrt{\beta_k}} \mathbf{d}_k + \sqrt{\beta_k} \hat{\mathbf{h}}_k \right) \left(\frac{1}{\sqrt{\beta_k}} \mathbf{d}_k + \sqrt{\beta_k} \hat{\mathbf{h}}_k \right)^H \right) \\ & \geq M - 1, \end{aligned} \quad (48)$$

where the last inequation holds true due to the fact that $\text{rank} \left(\left(\frac{1}{\sqrt{\beta_k}} \mathbf{d}_k + \sqrt{\beta_k} \hat{\mathbf{h}}_k \right) \left(\frac{1}{\sqrt{\beta_k}} \mathbf{d}_k + \sqrt{\beta_k} \hat{\mathbf{h}}_k \right)^H \right) = 1$. Recalling the complementary slackness conditions $\text{tr}(X_k \mathbf{Z}_k) = 0$ and $\text{rank}(X_k) \geq 1, \forall k$, we obtain $\text{rank}(\mathbf{Z}_k) = M - 1$ and $\text{rank}(X_k) = 1, \forall k$.

8.4 Appendix 4: Proof of Proposition 3

Firstly, we define $\mathbf{w}_k := \mathbf{w}_{k1} + j\mathbf{w}_{k2}$ (i.e., $\mathbf{w}_{k1} = \mathcal{R}(\mathbf{w}_k)$, and $\mathbf{w}_{k2} = \mathcal{I}(\mathbf{w}_k)$). Likewise, we have $\hat{\mathbf{h}}_k := \hat{\mathbf{h}}_{k1} + j\hat{\mathbf{h}}_{k2}$ and $\delta_k := \delta_{k1} + j\delta_{k2}$. Define $\mathbf{W}_{k,1} := [\mathbf{w}_{11}, \dots, \mathbf{w}_{k-1,1}, \mathbf{w}_{k+1,1}, \dots, \mathbf{w}_{K1}] \in \mathbb{R}^{M \times (K-1)}$ and $\mathbf{W}_{k,2} := [\mathbf{w}_{12}, \dots, \mathbf{w}_{k-1,2}, \mathbf{w}_{k+1,2}, \dots, \mathbf{w}_{K2}] \in \mathbb{R}^{M \times (K-1)}$. Let $\tilde{\mathbf{w}} := [\mathbf{w}_{11}^T, \mathbf{w}_{12}^T, \dots, \mathbf{w}_{K1}^T, \mathbf{w}_{K2}^T]^T$ and define K real $2K \times 2M$ matrices in terms of λ and $\tilde{\mathbf{w}}$ as

$$\mathbf{C}_k(\lambda, \tilde{\mathbf{w}}) := \begin{bmatrix} \frac{1}{\sqrt{\lambda \gamma_k}} \mathbf{w}_{k1} & \mathbf{W}_{k,1} & \mathbf{W}_{k,2} & \mathbf{0} \\ \frac{1}{\sqrt{\lambda \gamma_k}} \mathbf{w}_{k2} & \mathbf{W}_{k,2} & -\mathbf{W}_{k,1} & \mathbf{0} \end{bmatrix}^T, \forall k, \quad (49)$$

and K real $2K$ -dimension vectors in terms of λ and $\tilde{\mathbf{w}}$ as

$$\mathbf{c}_k(\lambda, \tilde{\mathbf{w}}) := \mathbf{C}_k(\lambda, \tilde{\mathbf{w}}) \begin{bmatrix} \hat{\mathbf{h}}_{k1} \\ \hat{\mathbf{h}}_{k2} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \sigma_k \end{bmatrix}, \forall k. \quad (50)$$

The robust SINR constraints (26) can be then written as

$$\mathbf{C}_k(\lambda, \tilde{\mathbf{w}}) \begin{bmatrix} \delta_{k1} \\ \delta_{k2} \end{bmatrix} + \mathbf{c}_k(\lambda, \tilde{\mathbf{w}}) \in \mathbb{L}^{2K+2}, \quad \forall \left\| \begin{bmatrix} \delta_{k1} \\ \delta_{k2} \end{bmatrix} \right\| \leq \epsilon_k. \quad (51)$$

Further define K real $2K \times (2M + 1)$ matrices as

$$\mathbf{B}_k(\lambda, \tilde{\mathbf{w}}) := [\mathbf{c}_k(\lambda, \tilde{\mathbf{w}}) \quad \epsilon_k \mathbf{C}_k(\lambda, \tilde{\mathbf{w}})],$$

and $\tilde{\delta}_k := [1 \quad \delta_{k1}^T \quad \delta_{k2}^T]^T \in \mathbb{R}^{2M+1}, \forall k$.

Note that $\mathbf{C}_k(\lambda, \tilde{\mathbf{w}})$ and $\mathbf{c}_k(\lambda, \tilde{\mathbf{w}})$ are affine with respect to \mathbf{w}_k , $\forall k$; so are $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$, $\forall k$. Then (51) becomes

$$\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})\tilde{\delta}_k \in \mathbb{L}^{2K}, \forall \tilde{\delta}_k \in \mathbb{L}^{2M+1}, \forall k. \quad (52)$$

Let

$$\tilde{\mathbf{D}}_m := \begin{bmatrix} \mathcal{R}(\mathbf{D}_m) & -\mathcal{I}(\mathbf{D}_m) \\ \mathcal{I}(\mathbf{D}_m) & \mathcal{R}(\mathbf{D}_m) \end{bmatrix} \in \mathbb{R}^{2M \times 2M}. \quad (53)$$

The per-antenna power constraints (7) are recast as SOCs:

$$\begin{bmatrix} \sqrt{\alpha P_m} \\ (\mathbf{I}_K \otimes \tilde{\mathbf{D}}_m)\tilde{\mathbf{w}} \end{bmatrix} \in \mathbb{L}^{2MK+1}, \quad \forall m, \quad (54)$$

where we use the fact that $\tilde{\mathbf{D}}_m = \tilde{\mathbf{D}}_m^{\frac{1}{2}}$.

Then (27) is transformed as the following real-valued optimization problem:

$$\begin{aligned} \hat{\alpha}^*(\lambda) &= \min_{\tilde{\mathbf{w}}, \alpha} \alpha \\ \text{s. t. } &\begin{bmatrix} \sqrt{\alpha P_m} \\ (\mathbf{I}_K \otimes \tilde{\mathbf{D}}_m)\tilde{\mathbf{w}} \end{bmatrix} \in \mathbb{L}^{2MK+1}, \quad \forall m, \\ &\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})\tilde{\delta}_k \in \mathbb{L}^{2K}, \forall \tilde{\delta}_k \in \mathbb{L}^{2M+1}, \forall k. \end{aligned} \quad (55)$$

In order to solve (55), let us first deal with the robust SOC constraints (i.e., the second group of constraints therein). For a given $\lambda > 0$, define set \mathcal{B}_k as

$$\mathcal{B}_k := \left\{ \mathbf{B}_k(\lambda, \tilde{\mathbf{w}}) | \mathbf{B}_k(\lambda, \tilde{\mathbf{w}})\mathbf{y}_k \in \mathbb{L}^{2K}, \forall \mathbf{y}_k \in \mathbb{L}^{2M+1} \right\}, \forall k.$$

It can be seen that \mathcal{B}_k contains linear maps taking \mathbb{L}^{2M+1} to \mathbb{L}^{2K} . In other words, \mathcal{B}_k is the set \mathcal{M} of LPMs such that $\mathcal{M}[\mathbb{L}^{2M+1}] \subset \mathbb{L}^{2K}$.

Lemma 6. The set \mathcal{B}_k consists of all Lorentz-positive matrices $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$, i.e.,

$$\mathbf{x}_k^T \mathbf{B}_k(\lambda, \tilde{\mathbf{w}})\mathbf{y}_k \geq 0, \forall \mathbf{x}_k \in \mathbb{L}^{2K}, \forall \mathbf{y}_k \in \mathbb{L}^{2M+1}. \quad (56)$$

Proof. It immediately follows from the fact that a Lorentz cone is self-dual [29].

Clearly, the set \mathcal{B}_k is a closed convex cone, and it thus has an LMI description [33]. With such a convex LMI description of \mathcal{B}_k , we are able to formulate the problem (55) into a standard SDP problem. For that purpose, we need to introduce some notations and facts as follows.

Let $\mathcal{S}(n)$ and $\mathcal{A}(n)$ be the spaces of real symmetric and skew-symmetric $n \times n$ matrices, respectively. $\mathcal{S}_+(n)$ is the cone of positive semidefinite matrices in $\mathcal{S}(n)$. Let $\mathcal{L}_{P,Q}$ denote the $[PQ(P+1)(Q+1)/4]$ -dimension linear space of biquadratic forms [36]:

$$\mathcal{L}_{P,Q} := \left\{ \begin{bmatrix} \mathbf{M}_{11} & \cdots & \mathbf{M}_{1P} \\ \vdots & \ddots & \vdots \\ \mathbf{M}_{P1} & \cdots & \mathbf{M}_{PP} \end{bmatrix} \mid \mathbf{M}_{pq} \in \mathcal{S}(Q) \right\}. \quad (57)$$

It is evident that $\mathcal{L}_{P,Q} \subset \mathcal{S}(PQ)$. The orthogonal complement of $\mathcal{L}_{P,Q}$ within $\mathcal{S}(PQ)$ is the $[PQ(P-1)(Q-1)/4]$ -dimension subspace $\mathcal{L}_{P,Q}^\perp$, where

$$\mathcal{L}_{P,Q}^\perp := \left\{ \begin{bmatrix} \mathbf{M}_{11} & \cdots & \mathbf{M}_{1P} \\ \vdots & \ddots & \vdots \\ \mathbf{M}_{P1} & \cdots & \mathbf{M}_{PP} \end{bmatrix} \mid \mathbf{M}_{pq} \in \mathcal{A}(Q) \right\}. \quad (58)$$

It is clear that $\mathcal{S}(PQ) = \mathcal{L}_{P,Q} \oplus \mathcal{L}_{P,Q}^\perp$, where \oplus denotes the direct sum of vector spaces. By these definitions, it can be seen that:

$$\mathcal{S}_+(P) \otimes \mathcal{S}_+(Q) \subseteq \mathcal{L}_{P,Q}, \quad \mathcal{A}(P) \otimes \mathcal{A}(Q) \subseteq \mathcal{L}_{P,Q}^\perp. \quad (59)$$

□

We recall the notation for $\hat{\mathbf{A}}(\cdot)$ in (28) and cite Theorem 5.6 in [33] as a lemma (see also [32], Lemma 3.2):

Lemma 7. Let $\min\{P, Q\} \geq 3$. Then a matrix $\mathbf{G} \in \mathbb{R}^{P \times Q}$ is Lorentz-positive, if and only if there exists $\mathbf{X} \in \mathcal{A}(P-1) \otimes \mathcal{A}(Q-1)$ such that

$$\hat{\mathbf{A}}(\mathbf{G}) + \mathbf{X} \succeq \mathbf{0}. \quad (60)$$

Based on (59) and Lemma 6, we can derive an equivalent condition for the Lorentz-positive $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$ in \mathcal{B}_k as follows.

Proposition 5. For a given $\lambda > 0$, $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$ is Lorentz positive if and only if there is $\mathbf{Z}_k \in \mathcal{L}_{2K-1, 2M}^\perp$ such that

$$\hat{\mathbf{A}}(\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})) + \mathbf{Z}_k \succeq \mathbf{0}, \quad \forall k. \quad (61)$$

Proof. The proof follows immediately from ([32] Proposition 3.3) and the observation in ([32] Equation (27)), and thus we omit it here.

We note that $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$ and $\hat{\mathbf{A}}(\mathbf{B}_k(\lambda, \tilde{\mathbf{w}}))$ (for a given $\lambda > 0$) are affine with respect to the design variable $\tilde{\mathbf{w}}$, and that $\mathcal{L}_{2K-1, 2M}^\perp$ is a linear subspace of symmetric matrices with skew-symmetric blocks. It follows that (61) is an implementable LMI description for $\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})$. Therefore, we obtain the following equivalent convex SDP reformulation of (55):

$$\begin{aligned} \hat{\alpha}^*(\lambda) &= \min_{\{\mathbf{Z}_k\}, \tilde{\mathbf{w}}, \alpha} \alpha \\ \text{s. t. } &\begin{bmatrix} \sqrt{\alpha P_m} \\ (\mathbf{I}_K \otimes \tilde{\mathbf{D}}_m)\tilde{\mathbf{w}} \end{bmatrix} \in \mathbb{L}^{2MK+1}, \forall m, \\ &\hat{\mathbf{A}}(\mathbf{B}_k(\lambda, \tilde{\mathbf{w}})) + \mathbf{Z}_k \succeq \mathbf{0}, \forall k, \\ &\mathbf{Z}_k \in \mathcal{L}_{2K-1, 2M}^\perp, \forall k. \end{aligned} \quad (62)$$

Now we complete the proof of Proposition 3. □

Endnotes

¹The proof can be found in our conference version [28].

²It can be ready to extend the results under the assumption of the spherical uncertainty region to the

ellipsoidal region case: $\mathcal{E}_k := \left\{ \hat{\mathbf{h}}_k + \delta_k \mid \delta_k^H \mathbf{B}_k \delta_k \leq \epsilon_k^2 \right\}$, where $\mathbf{B}_k, \forall k$, are positive definite matrices.

³Analogous to the separation property of convex set, any point \mathbf{z} outside a normal set can be separated from the normal set by a cone congruent to the nonnegative orthant. Thus, a normal set can be approximated as closely as desired by a nested sequence of polyblocks.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported in part by China Recruitment Program of Global Young Experts, the Program for New Century Excellent Talents in University, National Natural Science Foundation of China under Grant No. 61271223, and the National Science and Technology Major Project of the Ministry of Science and Technology of China under Grant No. 2012ZX03001013.

Author details

¹The Key Laboratory for Information Science of Electromagnetic Waves (MoE), the Department of Communication Science and Engineering, Fudan University, 220 Han Dan Road, Shanghai, China. ²Department of Mathematics, Hong Kong Baptist University, Hong Kong, China.

Received: 10 February 2015 Accepted: 21 July 2015

Published online: 16 August 2015

References

1. M Bengtsson, B Ottersten, in *Handbook of Antennas in Wireless Communications*, ed. by L Godara. Optimal and suboptimal transmit beamforming (CRC Press, New York, 2001)
2. M Schubert, H Boche, Solution of the multiuser downlink beamforming problem with individual SINR constraints. *IEEE Trans. Veh. Technol.* **53**(1), 18–28 (2004)
3. A Wiesel, YC Eldar, S Shamai, Linear precoding via conic optimization for fixed MIMO receivers. *IEEE Trans. Signal Process.* **54**(1), 61–176 (2006)
4. H Lee, I Sohn, D Kim, K-B Lee, in *Proc. IEEE ICC*. Generalized MMSE beamforming for downlink MIMO systems, Kyoto, Japan, 2011
5. L Qian, Y Zhang, J Huang, MAPEL: Achieving global optimality for a non-convex wireless power control problem. *IEEE Trans. Wireless Commun.* **8**(3), 1553–1563 (2009)
6. L Liu, R Zhang, K-C Chua, Achieving global optimality for weighted sum-rate maximization in the K-user Gaussian interference channel with multiple antennas. *IEEE Trans. Wireless Commun.* **11**(5), 1933–1945 (2012)
7. W Utschick, J Brehmer, Monotonic optimization framework for coordinated beamforming in multicell networks. *IEEE Trans. Signal Process.* **60**(4), 1899–1909 (2012)
8. C Xing, S Ma, Y Wu, Robust joint design of linear relay procedure and destination equalizer for dual-hop amplify-and-forward MIMO relay systems. *IEEE Trans. Signal Process.* **58**(4), 2273–2283 (2010)
9. X He, Y Wu, Probabilistic QoS constrained robust downlink multiuser MIMO transceiver design with arbitrary distributed channel uncertainty. *IEEE Trans. Wireless Commun.* **12**(12), 6292–6302 (2013)
10. K-Y Wang, AM-C So, T-H Chang, W-K Ma, C-Y Chi, Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization. *IEEE Trans. Signal Process.* **62**(21), 5690–5705 (2014)
11. M Shenouda, T N Davidson, Convex conic formulations of robust downlink precoder designs with quality of service constraints. *IEEE J. Sel. Topics Signal Process.* **1**(4), 714–724 (2007)
12. G Zheng, K-K Wong, T-S Ng, Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions. *EURASIP J. Adv. Signal Process.* **2008**, 1–15 (2008)
13. N Vucic, H Boche, Robust QoS-constrained optimization of downlink multiuser MISO systems. *IEEE Trans. Signal Process.* **57**(2), 714–725 (2009)
14. N Vucic, H Boche, S Shi, Robust transceiver optimization in downlink multiuser MIMO systems. *IEEE Trans. Signal Process.* **57**(9), 3576–3587 (2009)
15. E Song, Q Shi, M Sanjabi, R Sun, Z-Q Luo, Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight? *EURASIP J. Wireless Commun. Netw.* **2012**, 1687–1499 (2012)
16. G Zheng, K Wong, B Ottersten, Robust cognitive beamforming with bounded channel uncertainties. *IEEE Trans. Signal Process.* **57**(12), 4871–4881 (2009)
17. A Tajer, N Prasad, X Wang, Robust linear precoded design for multi-cell downlink transmission. *IEEE Trans. Signal Process.* **59**(1), 235–251 (2011)
18. E Bjornson, G Zhang, M Bengtsson, B Ottersten, Robust monotonic optimization framework for multicell MISO system. *IEEE Trans. Signal Process.* **60**(5), 2508–2523 (2012)
19. M Hanif, L-N Tran, A Tolli, M Juntti, S Glisic, Efficient solutions for weighted sum rate maximization in multicellular networks with channel uncertainties. *IEEE Trans. Signal Process.* **61**(22), 5659–5674 (2013)
20. W Yu, T Lan, Transmitter optimization for multi-antenna downlink with per-antenna power constraints. *IEEE Trans. Signal Process.* **55**(6), 2646–2660 (2007)
21. T Bogale, L Vandendorpe, Robust sum MSE optimization for downlink multiuser MIMO systems with arbitrary power constraint: Generalized duality approach. *IEEE Trans. Signal Process.* **60**(4), 1862–1875 (2012)
22. T Bogale, L Vandendorpe, Linear transceiver design for downlink multiuser MIMO systems: Downlink-interference duality. *IEEE Trans. Signal Process.* **61**(19), 4686–4700 (2013)
23. K Karakayali, R Yates, G Foschini, R Valenzuela, in *Proc. IEEE ISIT*. Optimum zero-forcing beamforming with per-antenna power constraints, Nice, France, 101–105 (2007)
24. P Ubaidulla, A Chockalingam, in *Proc. IEEE WCNC*. Robust THP transceiver designs for multiuser MIMO downlink, Budapest, Hungary, 2009
25. M Codreanu, A Tolli, M Juntti, M Latva-Aho, in *Proc. IEEE VTC-Spring*. MIMO downlink weighted sum rate maximization with power constraints per antenna groups, Dublin, Ireland, 2048–2052 (2007)
26. S Mohammed, E Larsson, Single-user beamforming in large-scale MISO system with per-antenna constant-envelope constraints: the doughnut channel. *IEEE Trans. Wireless Commun.* **11**(11), 3992–4005 (2012)
27. G Dartmann, X Gong, W Afzal, G Ascheid, On the duality of the max-min beamforming problem with per-antenna and per-antenna-array power constraints. *IEEE Trans. Veh. Technol.* **62**(2), 606–619 (2013)
28. F Wang, X Wang, Y Zhu, in *Proc. IEEE ICC*. Transmit beamforming for multiuser downlink with per-antenna power constraints, Sydney, Australia, 2014
29. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, Cambridge, UK, 2004)
30. I Polik, T Terlaky, A survey of the S-lemma. *SIAM Rev.* **49**(3), 371–418 (2007)
31. Z-Q Luo, W Ma, A So, Y Ye, S Zhang, Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Process. Mag.* **27**(3), 20–34 (2010)
32. Y Huang, D Palomar, S Zhang, Lorentz-positive maps and quadratic matrix inequalities with applications to robust MISO transmit beamforming. *IEEE Trans. Signal Process.* **61**(5), 1121–1130 (2013)
33. R Hildebrand, An LMI description for the cone of Lorentz-positive maps II. *Linear Multilinear Algebra.* **59**(7), 719–731 (2011)
34. H Tuy, Monotonic optimization: Problems and solution approaches. *SIAM J. Optim.* **11**(2), 464–494 (2000)
35. Z Fang, X Wang, X Yuan, Beamforming design for multiuser two-way relaying: a unified approach via max-min SINR. *IEEE Trans. Signal Process.* **61**(23), 5841–5852 (2013)
36. Z-Q Luo, J Strum, S Zhang, Multivariate nonnegative quadratic mappings. *SIAM J. Optim.* **14**(4), 1140–1162 (2004)